PHYLOGENETIC ANALYSIS OF ENOLASE GENE FAMILY IN DIFFERENT SPECIES

LIU, K. C. [1] – LIU, C.[1] – FAN, M. X.[2] – ZHANG, C. Y.[1] – CAI, F. C.[1] – LI, Q. Y.[1] – CHEN, M. M.[1] – JIANG, Y.[1] – HU, W. B.[1] – LIN, F.[1*]

[1]*College of Bioscience and Biotechnology, Shenyang Agricultural University, No. 120 Dongling Road, Shenyang, 110866 Liaoning, China*

[2]*Liaoning Key Laboratory of Urban Integrated Pest Management and Ecological Security, College of Life Science and Bioengineering, Shenyang University, Shenyang 110044, China*

*Corresponding author
e-mail: fenglinsn@126.com

Abstract. Enolase is involved in the glycolytic pathway of many organisms and is a key rate-limiting enzyme, which plays an important role in the growth and development of various species. In this study, the enolase gene sequences and amino acid sequences of more than 40 species were first downloaded, and phylogenetic analysis and molecular evolution studies were performed using bioinformatics methods. The phylogenetic tree was constructed to predict the conserved type of enolase evolution and the enolase evolution relationship of each species. It was found that the enolase of each species was very conservative during the evolution process. In addition to studying the evolutionary relationship of enolase of various species, the protein structure was predicted, and the secondary structure of enolase protein was predicted. The results showed that the secondary structure was very similar. The homology modeling method was used to predict the tertiary structure of enolase, and it was found that although there were slight differences, they were similar. In addition, statistical comparison analysis of the introns of each species enolase found that the gap between the number of introns and the locus was small, suggesting their close evolutionary relationship. It is predicted that a variety of plant and animal enolase catalytic kinase-specific phosphorylation sites, the results show that the kinase domain sequence contains three phosphorylation sites: serine (Ser), threonine (Thr) and tyrosine (Tyr). The amount of serine is the highest and the degree of similarity is high. The Ka/Ks ratio of most species is greater than 1. This indicates that enolase is mainly influenced by positive selection in evolution. Our results will help provide insights into the evolution of enolase and provide a theoretical basis for subsequent experiments.
Keywords: *enolase, glycolysis pathway, evolution, phosphorylation, intron*

Introduction

Enolase, originally discovered by Lohman and Meyerhof in 1934, is a rate-limiting enzyme in the glycolysis pathway. Enolase not only catalyzes the dehydration reaction from 2-phospho-D-glycerate (2-PGA) to phospho-enolpyruvate (PEP) in glycolysis metabolism, but also catalyzes the phospho-enolpyruvate hydratase reaction to convert PEP to 2-PGA (Lohman and Meyerhof, 1934; Armand et al., 2016) (*Fig. 1*). Further studies have demonstrated some divalent metal ions (Mg^{2+}, Mn^{2+}, Zn^{2+}, Ca^{2+}, Co^{2+} and Ni^{2+}) have activation effect on enolase activity, especially Mg^{2+} which is the strongest natural activator of enolase (Warburg et al., 1942; Mizuno et al., 2007). Enolase has long been considered as an ancient enzyme with only catalytic function. In addition to glycolytic pathways, enolase is also plays an important role in some physiological and pathological processes in organisms, such as ithe regulation of transcription, apoptosis, and cell differentiation (Straeten et al., 1991; Ucker et al., 2016).
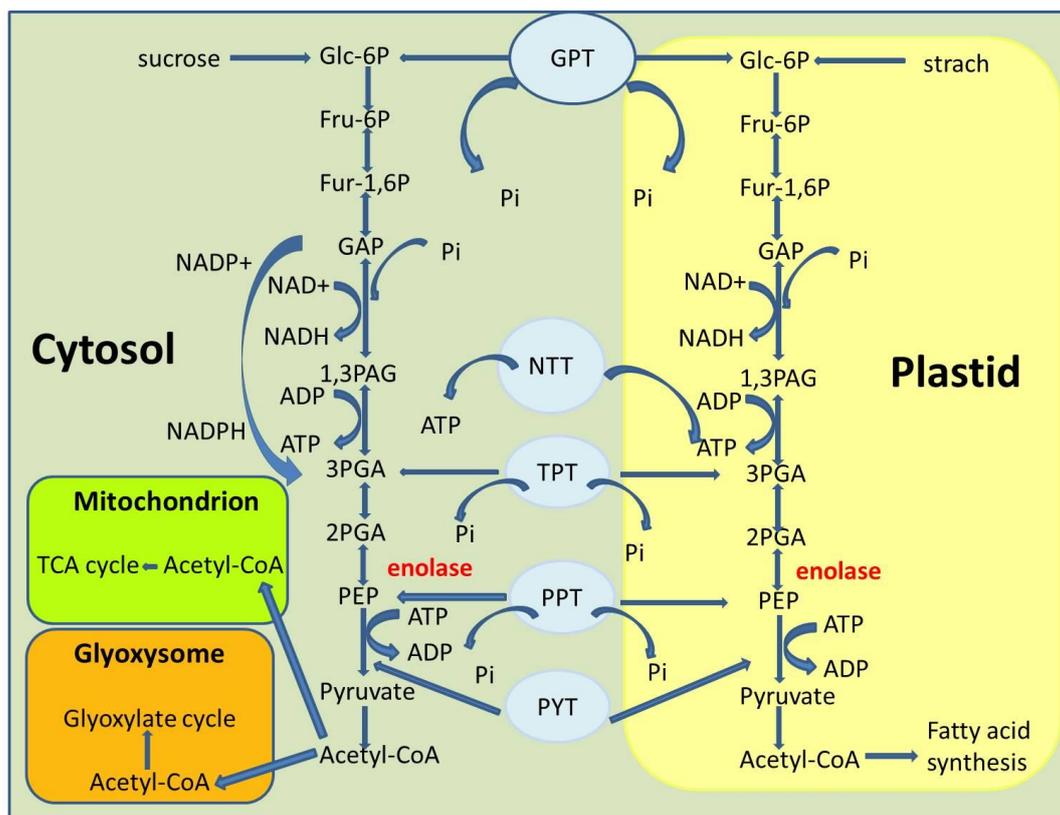
Figure 1. *Glycolysis pathway (PYT: pyruvate translocator; GPT: glucose phosphate/Pi translocator; PPT: phosphoenolpyruvate/Pi translocator; TPT: triose phosphate/Pi translocator; Fru-1, 6P: fructose 1,6 bisphosphate; Fru-6P: fructose 6 phosphate; GAP: glyceraldehyde-3-phosphate; Glc-6P: glucose 6 phosphate; PEP: phosphoenolpyruvate; 1,3-PGA: 1,3-bisphosphoglycerate;2-PGA: 2-phosphoglycerate; 3-PGA: 3-phosphoglycerate)*

Enolase occupies an important position in the fermentation metabolic pathway, especially in the glycolytic pathway. This enzyme is widely present in almost all the species. The gene sequences of enolases are highly conserved. The enolase is evolutionarily affiliated tothe enolase superfamily, which includes the MR (mandelate racemase) family, the MLE (muconate lactonizing enzyme) family, and the enolase family (Gerlt et al., 2012).

The crystal structures of enolase from some prokaryotes and eukaryotes have been resolved. The main structure of the enolase consists of two anti-parallel identical subunits. Each of these subunits contains a relatively long C-terminus and a relatively short N-terminal domain. The former domain is folded into a $\beta_2\alpha_2(\beta\alpha)_6$ topology, and the latter domain is folded into a $\beta_3\alpha_4$ topology (Lebioda et al., 1988). The activity center of enolase is located in the folded cavity of the C-terminal domain. Unlike the high degree of structural conservation and similarity, the aggregation state of enolase varies widely among different species.

Enolase is recognized as a multifunctional protein. In addition to the glycolysis function in metabolic pathway, enolase is also expressed as a receptor for plasminogen on the surface of many types of cells to ensure that plasminogen hydrolyzing activity is concentrated on the cell surface. This causes hydrolase cleavage of plasmin, which leads to the growth of tumor cell invasion and the associated inflammatory response

(Perconti et al., 2017). By removing the function of its plasminogen receptor and its catalytic function in glycolysis, enolase also exhibits other intracellular localization and cellular functions, all of which are well known for its glycolysis function. This suggests that it may play a role in cell recognition and invasion, vacuole formation and development, and transcriptional processes (Pal et al., 2009). In addition, differential expression of enolase is also thought to be associated with some pathophysiological states, such as cancer, Alzheimer's disease, rheumatoid arthritis. For example, a large number of studies have generally shown that the enzymatic enzyme's glycolysis function is significantly enhanced in tumor cells and is used as a marker for cancer progression (Song et al., 2014; Reev et al., 1986). Therefore, enolase is considered as a hallmark of many diseases. In plants studies about enolase have shown that the enolase gene expression and enzyme activity will change under high temperature, anaerobic, high salt, dry early, low temperature and other stress conditions (Barkla et al., 2009; Iida et al., 1985; Grondin et al., 2015). Under stress conditions, plants produce large amounts of free radicals that cause damage to the photosynthetic system. This stress response of enolase can provide energy for its repair. Enolase is a glycolytic enzyme that is consistent with its primary function in the cytoplasm. In the cytoplasm, enolase can be free, and other glycolytic enzymes in the cytoskeleton. Interaction occurs in combination (Subramanian et al., 2000). For example, pyruvate kinase, phosphoglycerate mutase and aldolase. The first two enzymes are closely adjacent to the enolase in the glycolytic pathway, and aldolase interacts with some of the cytoskeletal proteins (Aaronson et al., 1995). Studies have shown that α-enolase has a certain effect on hypoxia tolerance through non-glycolytic pathways. Under hypoxic conditions, the expression of α-enolase may be enhanced in cells. Protection under anaerobic conditions (Straeten et al., 1991).

We will carry out evolutionary cluster analysis of each species enolase, and analyze the protein secondary structure, tertiary structure, phosphorylation site, number of introns, and ka/ks value to prove the conservation of enolase. Through the research results of this paper, we can better analyze the evolutionary relationship between enolase in various species, and provide theoretical basis for subsequent research in molecular biology.

Materials and methods

Identification of enolase genes

In this study, the genes which encode enolase from 50 species were collected from databases and related articles, The relevant enolase sequences were obtained from phytozome database (available) online: (https://phytozome.jgi.doe.gov/pz/portal.html) and NCBI:(https://www.ncbi.nlm.nih.gov) results are merged and the extra sequences are removed. For genes with more than one predicted isotype, we only analyzed the longest peptide sequence for each gene.

Prediction of enolase structure

The amino acid sequences of enolase from 42 species were input into the SOPMA online software (https;//npsaprabi.ibcp.fr/cgi-bin) to predict the secondary structure of the enolase protein. And draw the obtained data into a table. Use SWISS-MODEL

(https://swissmodel.expasy.org/) online software to predict and build a tertiary structural model of enolase

The phylogram and selection of enolase

All enolase amino acid sequences in 50 species including maize, wheat, rice, Arabidopsis, mouse, orangutan were aligned by ClustalX v1.83 (Higgins et al., 1992). Use the default parameters. MEGA 6.0 (Tamura et al., 2011) using the NJ method was used to generate a phylogenetic tree. The enolase gene intron information for maize, wheat, rice, Arabidopsis, mice, orangutans, etc. is from Ensembl Genomes (Available online: http://ensemblgenomes.org/). The value of Ka/Ks was very important since it could estimate the changes of encoded amino acid sequences and gene functions, further estimate gene evolution. Each Ka and Ks value can be calculated by DnaSPv5.0 software (available online: http://www.ub.edu/dnasp/).

Phosphorylation analysis of enolase

KinasePhos (http://kinasephos.mbc.nctu.edu.tw/) was used to predict the kinase specific phosphorylation site in which enolase catalyzes. The software has high sensitivity and specificity, categorizing known phosphorylation sites by substrate sequences and their associated protein kinase classification.

The enolase gene intron information for 30 species including maize, rice, Arabidopsis, orangutan, and mouse is from Ensembl Genomes (available online at http://ensemblgenomes.org/).

Results

Conservative evolution of enolase

To investigate the gene evolutionary relationship, a total of 51 enolases from 50 species were aligned using MEGA software. The clustering analysis result showed that amino acid sequence levels in monkeys, bonobos, Sumatran orangutans, beluga whales, tilapia, rainbow trout, pigeons, ducks, birds, yeast, staphylococcus, staphylococcus, and streptococci. The enolase of 50 species was subjected to cluster analysis at the amino acid sequence level and merged with the system evolution, and the calibration parameters were repeated 1000 times (*Fig. 2*). The gymnosperms Araucaria and Selaginella are clustered together. Monocotyledonous corn, wheat, sorghum, and rice are clustered together, while most dicotyledonous plants such as peanuts, soybeans, pumpkins, and bitter gourds are clustered together, but peppers and tobacco. Gathering together as one, Arabidopsis and mulberry are clustered together. It may be that enolase has some differences in evolution among these species, which is worthy of our subsequent analysis. Enolase generally has the characteristics of closer evolution of the plant within the family. In mammals, mammals such as beluga whales, orangutans, monkeys, and horses are grouped together, rainbow trout in fish, tilapia are grouped together, mosquitoes, fruit flies, and silkworms are grouped together. Staphylococcus, streptococcus, and heatococcus are grouped together. Further analysis found that the similarity of several major crops such as wheat, corn, rice and sorghum was more than 54%, and the sequence similarity of beluga whales, horses, bonobos, red marmosets and Sumatran orangutans reached over 98%. This indicates that the enolase gene is evolutionarily conservative.
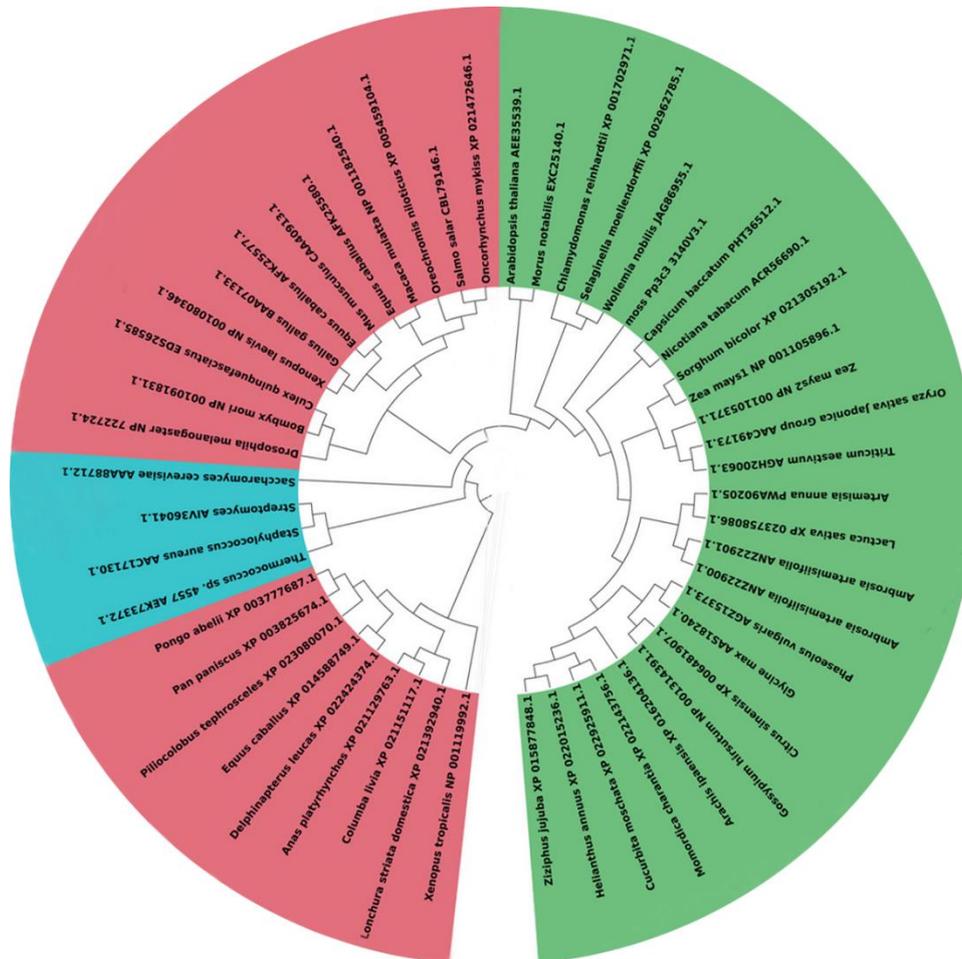
Figure 2. *Evolutionary analysis of the enolase amino acid sequence. (Note: Phylogenetic tree of enolase in 50 species including maize, rice, Arabidopsis, orangutan, and mouse. Green represents plants, pink represents animals, and blue represents fungi)*

Enolase protein structure analysis

The high-order structure formed by folding and winding of amino acid residues by van der Waals force, hydrogen bonding, etc. is called protein secondary structure, and mainly includes α-helix, β-sheet, random coil and extended chain. The secondary structure of 42 species of enolase protein was analyzed online using SOPMA software. The results showed that the enolase protein contained the above secondary structure, α-helix and random. Curl is the most frequently occurring structure, while the elongation chain and β-sheet appear less *(Schedule 1)*. The number of sputum-helix in plants is between 182 and 235, accounting for 39.96% to 46.53%. The number of "random" curls is between 143 and 171, accounting for 32.13% to 35.65%. In animals, the number of sputum-helix is between 184 and 292, accounting for 40.66% to 47%. The number of "random" curls is between 138 and 244, accounting for 31.8% to 40%. In the fungi, the number of α-helix is between 184 and 1945, accounting for 42.99%~44.7%. The number of "random" curls is between 136 and 143, accounting for 31.34% to 33.41%. The SWISS-MODEL uses the online tool SWISS-MODEL to model the tertiary structure of 42 species of enolase protein. After homologous modeling of 42 species of

enolase protein, it was found that the proteins are highly similar in spatial structure, and the proteins of each species have highly conserved structural features, so we constructed a tertiary structure of corn, orangutan, and thermococci enolase protein. The figures are representative of plants, animals, and bacteria (*Fig. 3a, b* and *c*). The results show that the tertiary structure of the enolase proteins of the three species is very similar. Both are bilaterally symmetric structures, and the positions of the active centers are also the same. And it can be seen intuitively that the α-helix, β-sheet, "random" curl, the number of extension chains and the sections are similar. However, the number of secondary structures of animal enolase is much higher than that of plants and fungi. Using Swiss-Pdb Viewer to analyze the results of homology modeling, the predicted dihedral angle of the protein residue is located in the yellow core region, indicating that its spatial structure is stable. PyMOL is used to process and analyze the modeling results, and the Ramachandran diagram is obtained. Ramachandran diagrams are mainly divided into permitted areas, maximum allowed areas, and disallowed areas. In general, the ratio of the amino acid of the model to the allowable region and the maximum allowable region is more than 90% of the total protein. The spatial structure conformation of the protein conforms to the rule of stereochemistry. As shown in *Figure 3d, e* and *f*, more than 90% of the dihedral angles of the enolase protein residues were detected in the allowable zone and the maximum allowable zone, indicating that it has a stable spatial conformation. In *Figure 3d, e* and *f* can be intuitively seen that the distribution of the secondary structure of the three representative species in the tertiary structure is concentrated in three main regions, and the representative species of the animal, in addition to the main three regions, are also distributed in other regions. This is different from the representatives of plants and fungi.
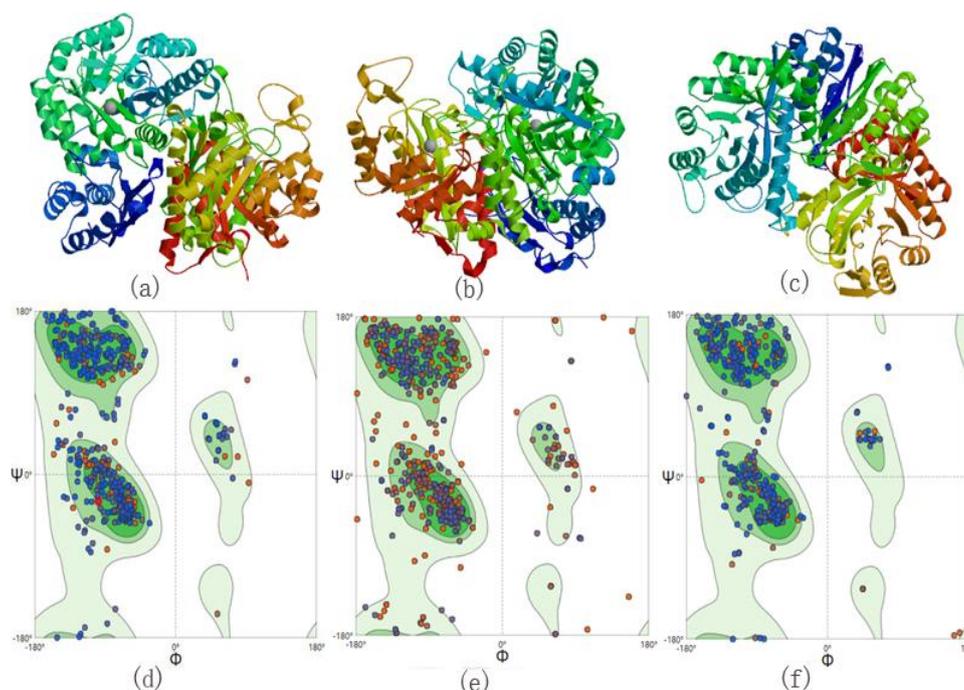

*Figure 3. Enolase tertiary structure diagram. **a** The tertiary structure of corn enolase, **b** the tertiary structure of the orangutan enolase, and **c** the tertiary structure of the thermococcal enolase. **d** A PyMOL map representing the tertiary structure of corn enolase, **e** a PyMOL map representing the tertiary structure of the orangutan enolase, and **f** a PyMOL map representing the tertiary structure of the thermococcal enolase*

Enolase phosphorylation site analysis

Phosphorylation of proteins is one of the important covalent modifications in organisms. About one-third of proteins are hyperphosphorylated during mammalian cell life cycle; 5% of genes in vertebrate genomes The encoded protein is a protein kinase and phosphoenzyme involved in the phosphorylation and dephosphorylation processes. Plants are grown in the soil and cannot move away from biological or abiotic stresses like animals. Therefore, plant cells usually use receptors on the cell surface to sense environmental changes, transmit signals through various signaling pathways in the body, and finally respond appropriately to various external stimuli (Morris and Walker, 2003). The protein phosphorylation sites of eukaryotic cells mainly occur on the hydroxyl groups of the serine (Ser), threonine (Thr) and tyrosine (Tyr) residues, and different protein kinases can recognize and modify the difference of different proteins Site. Kinase Phos (http://kinasephos.mbc.nctu.edu.tw/) was used to predict the kinase-specific phosphorylation sites that 14 plants and 11 animal enolase catalyze. The software is highly sensitive and specific, categorizing known phosphorylation sites by substrate sequences and their associated protein kinase classification. According to the prediction results of phosphorylation sites (*Fig. 4a* and *b*), the amino acid sites with a predicted score of 0.5 or more were recorded. The results showed that the kinase domain sequence contains three phosphorylation sites: serine, serine (Thr) and tyrosine (Tyr). Among them, the potential of serine phosphorylation sites is the highest in animals. Plants have the highest levels of potential serine phosphorylation sites and tyrosine phosphorylation sites. There are 6 phosphorylation sites in rice, wheat and jujube. Maize, citrus, kidney bean, sorghum and peanut contain 5 phosphorylation sites, and corn, rice, wheat, sorghum, jujube and citrus contain Thr47 and Thr59 phosphate. Chemical site. Peanut, kidney bean, soybean, and bitter gourd all contain Thr46 and Thr58 phosphorylation sites, and tobacco contains Thr46 phosphorylation site. Rice, wheat and jujube contain Ser277 phosphorylation sites, citrus, peanuts, and Ser276 phosphorylation sites, and bitter melon and soybean contain Ser275 phosphorylation sites. Kidney beans, soybeans, and bitter gourd contain Ser382 phosphorylation sites, citrus, peanuts, and Ser383 phosphorylation sites, and corn, sorghum, and jujube contain Ser384 phosphorylation sites. Sumatra orangutans, horses, monkeys, and beluga contain Ser289, Ser527, and Thr189 phosphorylation sites.

Enolase intron number analysis

An intron is a non-coding sequence of a break gene that can be transcribed but is cleaved off during mRNA processing, so there is no intron coding sequence on the mature mRNA. In eukaryotes, most genes are interrupted by one or more introns, which may contain "old codes", which are part of the gene that loses its function during evolution. It may be a key event in the conservative evolution of gene structure. Because the intron is meaningless to the structure of the translation product and is not subject to the pressure of natural selection, it has more mutations than the exon accumulation. In addition, there is growing evidence that loss or increase in introns may be an important factor in the progressive diversity and complexity of genes. We downloaded the enolase genes from 30 species and counted the number of introns. The results are shown in *Figure 5*. Among them, the main crops of maize, rice, sorghum, soybean, and wheat enolase genes had 16 introns. The number of animal introns is mainly concentrated in 11, 13, and 14.
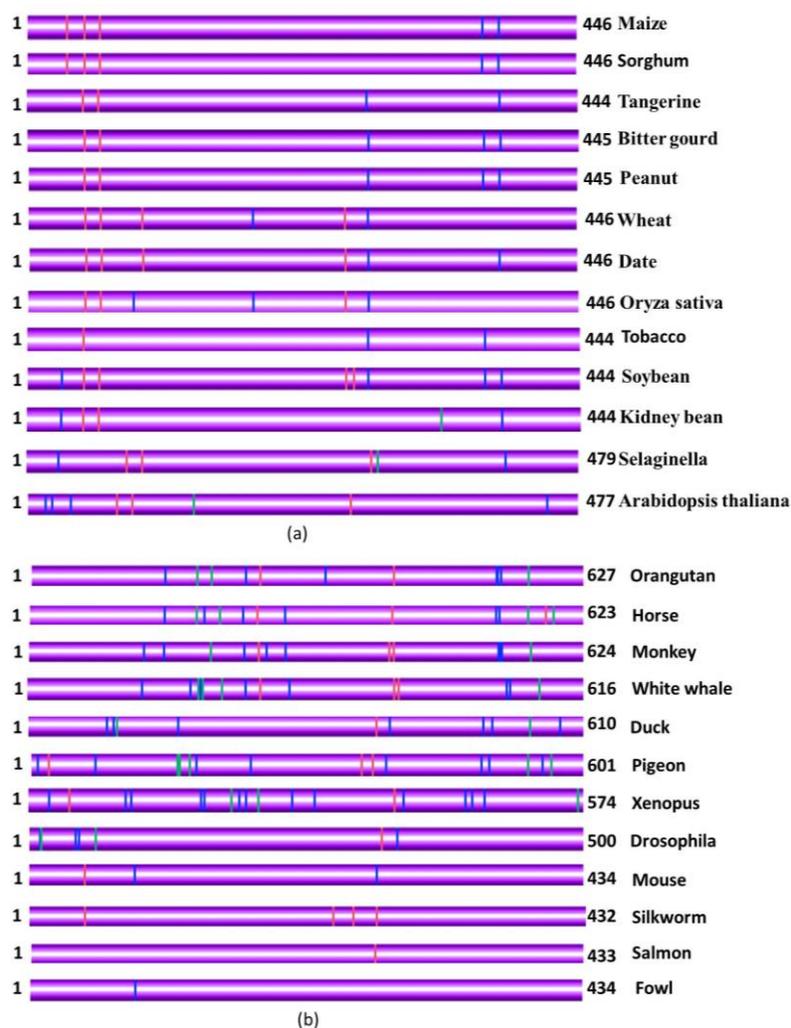
*Figure 4. Enolase phosphorylation site analysis. (Note: Phosphorylation sites of enolase in 24 species such as maize, rice, Arabidopsis, orangutan, and mouse. Red represents serine, green represents threonine, and blue represents tyrosine. **a** The position and number of phosphorylation sites of plant enolase, and **b** the location and number of phosphorylation sites of animal enolase)*

The evolution selection of enolase

In genetics, Ka/Ks or dN/dS represents the ratio between the heterosexual substitution (Ka) and the synonymous substitution (Ks). This ratio can be used to determine if there is a selective pressure on this protein-coding gene. Nucleotide variants that do not result in amino acid changes are referred to as synonymous mutations, and vice versa. It is generally believed that synonymous mutations are not naturally selected, while non-synonymous mutations are naturally selected. Because synonymous substitutions cause amino acid changes that may alter the conformation and function of the protein, and thus cause adaptive changes that result in advantages or disadvantages of natural selection (generally disadvantages). Non-synonymous substitutions do not alter the composition of the protein and are therefore unaffected by natural selection, so Ks can reflect the background base substitution rate of the evolutionary process. The ratio of Ka/Ks can indicate which choice the gene is. In

evolutionary analysis, it is meaningful to understand the rate at which synonymous and non-synonymous mutations occur. Commonly used parameters are the following: synonymous mutation frequency (Ks), non-synonymous mutation frequency (Ka), ratio of non-synonymous mutation rate to synonymous mutation rate (Ka/Ks). If Ka/Ks > 1, then a positive selection effect is considered. If Ka / Ks = 1, then a neutral choice is considered to exist. If Ka/Ks < 1, then a purification selection effect is considered (Baumgarten et al., 2003). The frequency of enolase Ka/Ks in plants is mainly between 1.2 and 1.4, and all Ka/Ks are almost > 1 (*Fig. 6a* and *b*). The frequency of enolase Ka/Ks in animals is mainly between 1.0 and 1.2. Although most of the Ka/Ks are > 1, there are certain ratios between 0.8 and 1.0 indicating (*Fig. 6c* and *d*) that the enolase gene is mainly influenced by the positive selection effect in evolution. The enolase in plants is almost always influenced by the positive selection effect in evolution. The enolase in animals is influenced by the positive selection effect in evolution, but some are also affected by negative selection.
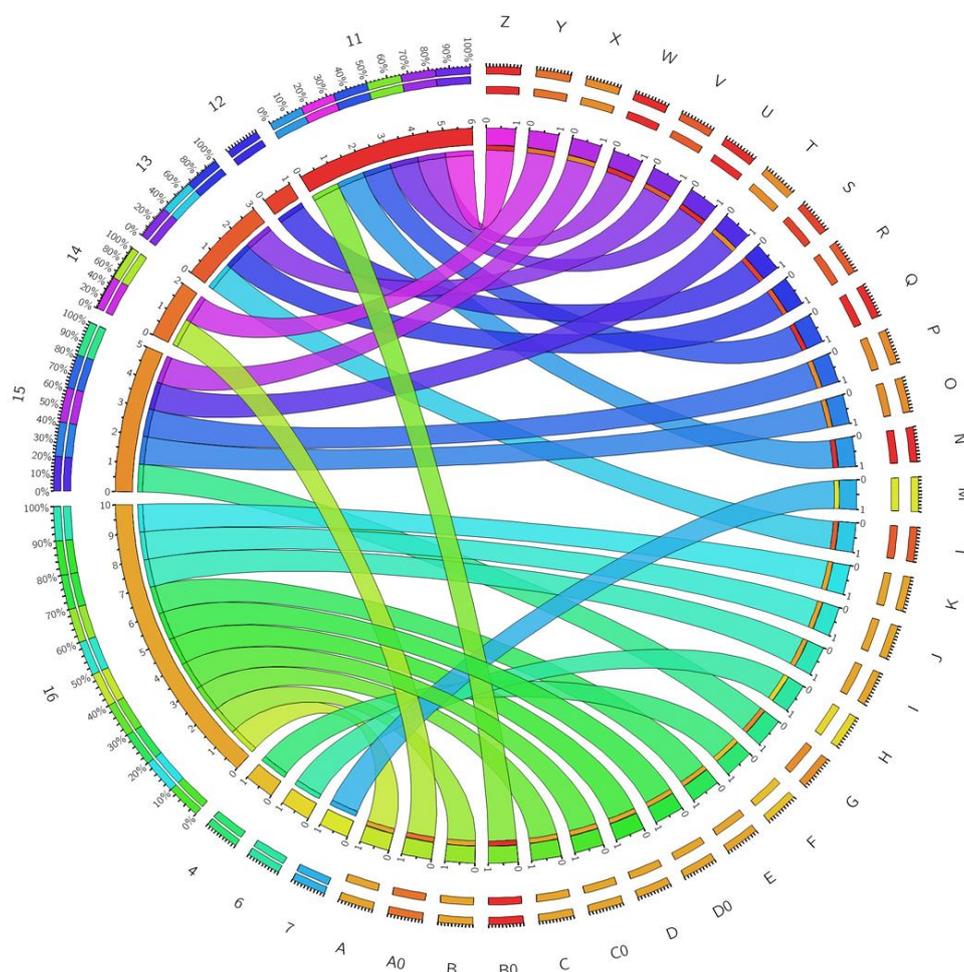

Figure 5. *Enolase intron analysis. (Note: Circos map of the enolase gene in 30 species including maize, rice and Arabidopsis. The number on the left represents the number of introns, the letter on the right is the code of the species. A: corn, B: corn, c: soybean, D: sorghum, E: peanut, F: yeast, G: cotton, H: Arabidopsis, I: rice, J: wheat, K: tobacco, L: orangutan, M: silkworm, N: Xenopus laevis, O: citrus, P: pigeon, Q: squid, R: monkey, S: Selaginella, T: bitter gourd, U: tilapia, V: duck, W: horse, X: Pumpkin, Y: Sumatran orangutan, Z: Lettuce, A0: Sunflower, B0: Mouse, CO: Jujube, D0: Sesame)*
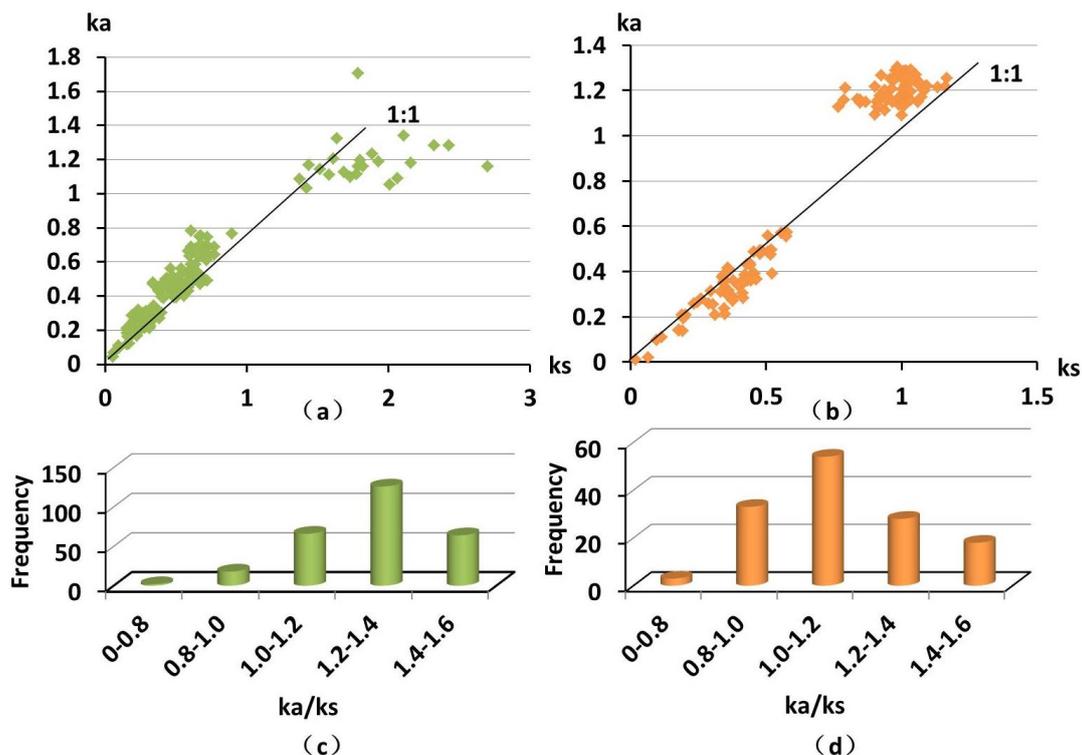
Figure 6. *The distribution and frequency of Ka/Ks values. (Note: The orange and green colors represent the animal and plant respectively.* ***a, b*** *The Ka/Ks distribution. The line in black is the boundary of Ka/Ks = 1.* ***c, d*** *The frequency distribution of Ka/Ks)*

Discussion

As one of the key enzymes in the glycolytic pathway, Enolase catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate during glycolysis, which is essential for cell anabolism and energy production (Pancholi et al., 2001). The enolase gene of multiple species was studied, and the genetic evolution relationship, molecular structure and protein function expressed by these genes were analyzed based on bioinformatics methods. It was found that the enolase candidate amino acid sequences of 50 species have high similarity, some even higher than 99%, indicating that they are highly conserved during evolution. From the constructed phylogenetic tree (*Fig. 2*), it can be seen that the gymnosperms are clustered into one, the monocotyledons are clustered into one, and most of the dicotyledons are also clustered into one, but the dicotyledonous plants of pepper and tobacco are gathered together as one. The Arabidopsis and mulberry trees are clustered together and do not cluster with other dicotyledonous plants, probably due to the evolutionary evolution of the enolase genes in these four plants to the rest of the plants. However, in general, enolase has the characteristics of closer evolution of the plant. Animals are divided into two branches, each containing mammals, one of which includes fruit fly, mouse, horse, fish and fungal branches are relatively close, presumably due to the more conservative enolase gene in each species There is little difference within the long process, and only slight changes in the long evolutionary process, so the molecular evolution trend of animal enolase is inconsistent with the evolutionary trend of species. This situation has also appeared in the evolution of other genes. The evolutionary distance of the miR166 gene family is

not related to the genetic relationship between the species itself (Xu-Yao et al., 2014) speculate that the gene may evolve in different ways and speeds. The similarity of the enolase amino acid sequence of each species is very high. The similarity between the bean and soybean, sorghum and corn in the plant reaches 100%, the similarity between bitter gourd and pumpkin reaches 99%, and the beluga whale, horse, bonobo, red marmoset in the animal. The sequence similarity of Sumatra orangutans reached over 98%, and the similarity of other animals was above 80%. This indicates that enolase is highly conserved in evolution. Next, we analyze the introns of the enolase gene of each species. Generally, the number of introns is similar, and the intron position and phase between genes with closer evolution are more similar. The number of introns of the enolase gene in most species is concentrated in 11-16 (*Fig. 3*). It shows that the evolutionary relationship of enolase is very close in different species. It is worth noting and discussing that only six of the introns in the Arabidopsis enolase gene are the same as the other species. Enolase is not expressed in the leaves of Arabidopsis, and if silenced, it affects the development of Arabidopsis sporozoites and gametophytes (Prabhakar et al., 2010) and the absence of enolase also affects the root development of Arabidopsis (Prabhakar et al., 2009). The Arabidopsis enolase structure has a pair of disulfide bonds formed by Cys residues, which can undergo redox regulation. The enolase in corn, tomato and ramie does not contain this pair of Cys residues. It seems possible that enolase is redox-regulated by a Cytosolic thioredoxin system in a limited number of plant species including ice plant and Arabidopsis (Anderson et al., 1998). We can later study what kind of particularity the enolase has in its structure and evolution with other plants. The enolase phosphorylation site of each species was predicted to have the highest number of serines and high similarity (*Fig. 4*). The eight phosphorylation sites of Thr46, Thr47, Thr58, Ser275, Ser276, Ser277, Ser382 and Ser383 appeared most frequently in plants. The three phosphorylation sites of Ser289, Ser527 and Thr189 appeared most frequently in animals. The phosphorylation sites of plant enolase are more conservative in evolution than animals, both in number and in location. The similarity of phosphorylation sites in animals is not very high. There are only three phosphorylation sites at high frequencies, and the phosphorylation sites of similar genetically related species are still very similar, such as mammalian orangutans and beluga whales, horse. Most protein interactions are carried out through the activation of phosphating sites, which provides a basis for the subsequent interaction of enolase with its upstream and downstream related proteins. Secondary structure analysis of 42 species of enolase showed that α-helix and "random" crimp were conserved structures, while α-helix and "random" curls were the most numerous, so α-helix and "Random" curl plays an important role in the formation of advanced structures of enolase. The prediction of secondary structure indicates that the secondary structure of protein determines its tertiary structure, which is closely related to the function of protein expression. Based on the similarity of secondary structure of enolase of each species, its functional similarity is predicted. Very conservative, the difference is not big. It can be seen from the (*Fig. 5*) that the tertiary structure of the enolase of each species is also very similar. It is generally accepted that a case where a nucleotide variation does not cause an amino acid change is called a synonymous mutation (Ks), and a phenomenon in which a nucleotide mutation also causes an amino acid change is called a non-synonymous mutation (Ka). Usually, under natural selection, synonymous mutations are not affected, but synonymous mutations are affected by natural selection (Yang et al., 2006). Their ratio Ka/Ks is a molecular evolutionary parameter that weighs

the pressure of selection to determine whether there is a selective pressure on the gene-encoded gene (Shiu et al., 2004). As shown in (*Fig. 6*), it is not that there are positive selection sites for most enolase genes in plants or animals. Some of them have a positive selection site, but their Ka/Ks values are close to 1, indicating positive selection and natural selection for enolase. Genetic evolution affects key roles. Such genes are genes that are rapidly evolving in the near future and are very important for the evolution of species. In summary, we conducted a comparative study on the enolase genes of different species by bioinformatics methods, and found that the enolase gene evolution is very conservative and similar in function. Since enolase is a key rate-limiting enzyme in the glycolytic pathway, the results of this study can provide an analytical basis for subsequent molecular experiments, and are highly versatile in terms of increased yield, resistance to stress, cancer intervention, and treatment in GM crops prospects (Altenberg et al., 2004; Pu et al., 2007).

REFERENCES

[1] Aaronson, R. M. (1995): Non-neuronal enolase is an endothelial hypoxic stress protein. – Journal of Biological Chemistry 270(46): 27752-7.

[2] Altenberg, B., Greulich, K. O. (2004): Genes of glycolysis are ubiquitously overexpressed in 24 cancer classes. – Genomics 84(6): 1014-1020.

[3] Anderson, L. E., Li, A. D., Stevens, F. J. (1998): The enolases of ice plant and arabidopsis contain a potential disulphide and are redox sensitive. – Phytochemistry 47(5): 707-13.

[4] Armand, D. A., Flores-Tornero, M., Rosa-Telléz, S., Muñoz-Bertomeu, J., Segura, J., Ros, R. (2016): The specific role of plastidial glycolysis in photosynthetic and heterotrophic cells under scrutiny through the study of glyceraldehyde-3-phosphate dehydrogenase. – Plant Signal. Behavior 11: e1128614.

[5] Barkla, B. J., Vera-Estrella, R., Hernandez-Coronado, M., Pantoja, O. (2009): Quantitative proteomics of the tonoplast reveals a role for glycolytic enzymes in salt tolerance. – The Plant Cell 21(12): 4044-4058.

[6] Baumgarten, A., Cannon, S., Sprangler, R., May, G. (2003): Genome-level evolution of resistance genes in Arabidopsis thaliana. – Genetics 165(1): 309-319.

[7] Gerlt, J. A., Babbitt, P. C., Jacobson, M. P., Almo, S. C. (2012): Divergent evolution in enolase superfamily: strategies for assigning functions. – Journal of Biological Chemistry 287(1): 29-34.

[8] Grondin, M., Chow-Shi-Yée, M., Ouellet, F., Averill-Bates, D. A. (2015): Wheat enolase demonstrates potential as a non-toxic cryopreservation agent for liver and pancreatic cells. – Biotechnology Journal 10(5): 801-810.

[9] Higgins, D. G., Bleasby, A. J., Fuchs, R. (1992): Clustal v: improved software for multiple sequence alignment. – Comput Appl Biosci 8(2): 189-191.

[10] Iida, H., Yahara, I. (1985): Yeast heat-shock protein of mr 48,000 is an isoprotein of enolase. – Nature 315(6021): 688-690.

[11] Lebioda, L., Stec, B. (1988): Crystal structure of enolase indicates that enolase and pyruvate kinase evolved from a common ancestor. – Nature 333(6174): 683-686.

[12] Lohman, K., Meyerhof, O. (1934): Über die enzymatische Umwandlung von Phosphoglyzerinsäure in Brenztraubensäure und Phosphorsäure (Enzymatic

transformation of phosphoglyceric acid into pyruvic and phosphoric acid). – Biochem. Z. 273: 60-72.

[13] Pal Bhowmick, I., Kumar, N., Sharma, S., Coppens, I., Jarori, G. K. (2009): Plasmodium falciparumenolase: stage-specific expression and sub-cellular localization. – Malaria Journal 8(1): 179.

[14] Pancholi, V. (2001): Multifunctional α-enolase: its role in diseases. – Cellular & Molecular Life Sciences Cmls 58(7): 902-920.

[15] Perconti, G., Maranto, C., Romancino, D. P., Rubino, P., Feo, S., Bongiovanni, A., et al. (2017): Pro-invasive stimuli and the interacting protein hsp70 favour the route of alpha-enolase to the cell surface. – Scientific Reports 7(1): 3841.

[16] Piast, M., Kustrzeba-Wójcicka, I., Matusiewicz, M., Banaś, T. (2005): Molecular evolution of enolase. – Acta Biochimica Polonica 52(2): 507-13.

[17] Prabhakar, V., Löttgert, T., Gigolashvili, T., Bell, K., Ulf-Ingo Flügge, Häusler, R. E. (2009): Molecular and functional characterization of the plastid-localized phosphoenolpyruvate enolase (eno1) from Arabidopsis thaliana. – FEBS letters 583(6): 983-991.

[18] Prabhakar, V., Löttgert, T., Geimer, S., Dörmann, P., Krüger, S., Vijayakumar, V., et al. (2010): Phosphoenolpyruvate provision to plastids is essential for gametophyte and sporophyte development in Arabidopsis thaliana. – Plant Cell 22(8): 2594-2617.

[19] Pu, L. J., Huang, S. M., Liu, F. (2007): Effects of losartan on expressions of plasminogen activator and plasminogen activator inhibitor-1 of rat proximal tubular epithelial cells cultured with high glucose. – Medical science edition 38(5): 813.

[20] Reeve, J. G., Stewart, J., Watson, J. V., Wulfrank, D., Twentyman, P. R., Bleehen, N. M. (1986): Neuron specific enolase expression in carcinoma of the lung. – British Journal of Cancer 53(4): 519-528.

[21] Song, Y., Luo, Q., Long, H., Hu, Z., Que, T., Xi'an Zhang, et al. (2014): Erratum to: alpha-enolase as a potential cancer prognostic marker promotes cell growth, migration, and invasion in glioma. – Molecular Cancer 13(1): 235.

[22] Straeten, D. V. D., Rodrigues-Pousada, R. A., Montagu, G. M. V. (1991): Plant enolase: gene structure, expression, and evolution. – The Plant Cell 3(7): 719-735.

[23] Subramanian, A. (2000): Structural analysis of alpha -enolase. Mapping the functional domains involved in down-regulation of the c-myc protooncogene. – Journal of Biological Chemistry 275(8): 5958-5965.

[24] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011): MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. – Mol. Biol. Evol. 28: 2731-2739.

[25] Ucker, S. D. (2016): Exploiting death: apoptotic immunity in microbial pathogenesis. – Cell Death and Differentiation 23: 990-996.

[26] Walker, T., Morris, J., Threlfall, R., Main, G. (2003): Analysis of wine components in cynthiana and syrah wines. – Journal of Agricultural & Food Chemistry 51(6): 1543.

[27] Warburg, O., Christian, W. (1942): Isolation and crystallization of enolase. – Biochim. Z. 310: 384-421.

[28] Xu-Yao, Z., Si-Yun, C., Lei, Z., et al. (2003): Evolution of MIR166 gene family in land plants. – Plant Diversity and Resources 36(3): 331-341.

APPENDIX

	Alpha helix (%)		Beta turn (%)		Random coil (%)		Extended strand (%)	
Glycine max	191	43.02%	34	7.66%	149	33.56%	70	15.77%
Citrus sinens	193	43.37%	40	8.99%	143	32.13%	69	15.51%
Sorghum bicolor	185	41.48%	33	7.40%	159	35.65%	69	15.47%
Arach ipaens	191	42.92%	36	8.09%	149	33.48%	69	15.51%
Saccharomyces cereviae	190	43.48%	36	8.24%	147	33.64%	64	14.65%
Selaginella moellendorffii	222	46.35%	31	6.47%	155	32.36%	71	14.82%
Momordica charantia	193	43.47%	40	9.01%	143	32.21%	68	15.32%
Capsicum baccatum	235	45.11%	33	6.33%	171	32.82%	82	15.74%
Chlamydomonas reinhardtii	218	45.70%	33	6.92%	159	33.33%	67	14.05%
Gossypium hirsutum	190	42.70%	37	8.31%	150	33.71%	68	15.28%
Cucurbita moschata	195	43.92%	39	8.78%	140	31.53%	70	15.77%
Arabidops thaliana	205	42.98%	37	7.76%	154	32.29%	81	16.98%
Artemia annua	191	43.02%	39	8.78%	150	33.78%	64	14.41%
Morus notabil	215	39.96%	37	6.88%	201	37.36%	85	15.80%
Oryza sativa	192	43.05%	35	7.85%	152	34.08%	67	15.02%
Ambrosia artemiifolia	184	42.69%	35	8.12%	147	34.11%	65	15.08%
Culex quinquefasciatus	182	42.03%	30	6.93%	153	35.33%	68	15.70%
Lactuca sativa	186	41.89%	36	8.11%	148	33.33%	74	16.67%
Triticum aestivum	197	44.17%	33	7.40%	154	34.53%	62	13.90%
Nicotiana tabacum	189	42.57%	38	8.56%	151	34.01%	66	14.86%
Zea mays	192	43.05%	36	8.07%	146	32.74%	72	16.14%
Moss	196	44.04%	36	8.09%	149	33.48%	64	14.38%
Ziziphus jujuba	191	42.83%	35	7.85%	153	34.30%	67	15.02%
Pan pancus	292	46.57%	26	4.15%	241	38.44%	68	15.02%
Delphinapterus leucas	288	46.75%	34	5.52%	230	37.34%	64	10.39%
Xenopustropical	273	47.56%	35	6.10%	211	36.76%	55	9.58%
Pongo abelii	288	46.08%	35	5.60%	238	38.08%	64	10.24%
Equus caballus	184	42.40%	32	7.37%	153	35.25%	65	14.98%
Piliocolobus tephrosceles	274	43.91%	45	7.21%	228	36.54%	77	12.34%
Mus musculus	184	42.40%	31	7.14%	147	33.87%	72	16.59%
Lonchura striata domestica	234	43.82%	41	7.68%	190	35.58%	69	12.92%
Anas platyrhynchos	248	40.66%	42	6.89%	244	40.00%	76	12.46%
Gallus gallus	190	43.78%	33	7.60%	143	32.95%	68	15.67%
Columba livia	263	43.76%	30	4.99%	245	40.77%	63	10.48%
Oreochrom niloticus	192	44.24%	35	8.06%	138	31.80%	69	15.90%
Salmo salar	189	43.75%	32	7.41%	140	32.41%	71	16.44%
Drosophila melanogaster	200	40.00%	35	7.00%	181	36.20%	84	16.80%
Bombyx mori	189	43.65%	35	8.08%	145	33.49%	64	14.78%
Streptomyces	184	42.99%	34	7.94%	143	33.41%	67	15.65%
Staphylococcus aureus	194	44.70%	38	8.76%	136	31.34%	66	15.21%
Thermococcus	187	43.49%	34	7.91%	140	32.56%	69	16.05%