



ISSN 1008-9209

CODEN ZXSKFJ

JOURNAL OF  
ZHEJIANG UNIVERSITY

AGRICULTURE AND LIFE SCIENCES

# 浙江大学学报

农业与生命科学版

ISSN 1008-9209



9 771008 920140

**4/2014**

Vol.40 No.4 July 2014

中国·杭州 HANGZHOU, CHINA  
Zhejiang Daxue Xuebao (Nongye yu Shengming Kexue Ban)

# 《浙江大学学报(农业与生命科学版)》编委会

## Editorial Board of Journal of Zhejiang University (Agriculture & Life Sciences)

### 主编 Editor-in-Chief

朱 军 (Zhu Jun, Zhejiang University, jzhu@zju.edu.cn)

### 副主编 Associate Editor-in-Chief

吴 平 (Wu Ping, Zhejiang University)

### 国外编委 Foreign Members of the Editorial Board

Chrisman Cheryl L., University of Florida, USA

Kohanawa Masashi, Hokkaido University, Japan

Lai Chaoqiang, Tufts University, USA

May Sean, University of Nottingham, UK

Qin Jianguang, Flinders University, Australia

Song Guoqing, Michigan State University, USA

Stafford James L., University of Alberta, Canada

Tang Lie, Iowa State University, USA

Whelan James M., University of Western Australia, Australia

Xie Jiahua, North Carolina Central University, USA

Zhao Fengqi, University of Vermont, USA

Zhu Qianhao, CSIRO Plant Industry, Australia

### 国内编委 Domestic Members of the Editorial Board

包劲松 (Bao Jinsong, Zhejiang University)

刁宏燕 (Diao Hongyan, Zhejiang University)

方维焕 (Fang Weihuan, Zhejiang University)

管敏鑫 (Guan Minxin, Zhejiang University)

何 勇 (He Yong, Zhejiang University)

胡元亮 (Hu Yuanliang, Nanjing Agricultural University)

李民赞 (Li Minzan, China Agricultural University)

刘建新 (Liu Jianxin, Zhejiang University)

刘向东 (Liu Xiangdong, Zhejiang Sci-Tech University)

缪云根 (Miao Yungen, Zhejiang University)

汪俏梅 (Wang Qiaomei, Zhejiang University)

吴信忠 (Wu Xinzong, Zhejiang University)

叶兴乾 (Ye Xingqian, Zhejiang University)

章明奎 (Zhang Mingkui, Zhejiang University)

赵宇华 (Zhao Yuhua, Zhejiang University)

周东美 (Zhou Dongmei, Chinese Academy of Sciences)

陈 铭 (Chen Ming, Zhejiang University)

樊龙江 (Fan Longjiang, Zhejiang University)

傅 衍 (Fu Yan, Zhejiang University)

郭仰东 (Guo Yangdong, China Agricultural University)

胡松华 (Hu Songhua, Zhejiang University)

李 铎 (Li Duo, Zhejiang University)

梁月荣 (Liang Yuerong, Zhejiang University)

刘维屏 (Liu Weiping, Zhejiang University)

马忠华 (Ma Zhonghua, Zhejiang University)

王 俊 (Wang Jun, Zhejiang University)

王 岩 (Wang Yan, Zhejiang University)

杨肖娥 (Yang Xiaoe, Zhejiang University)

张国平 (Zhang Guoping, Zhejiang University)

张子丁 (Zhang Ziding, China Agricultural University)

郑 平 (Zheng Ping, Zhejiang University)

朱伟云 (Zhu Weiyun, Nanjing Agricultural University)

### 执行编辑 Managing Editor

胡清华 (Hu Qinghua, zdxbsb@zju.edu.cn)

目 次 Contents

本刊编委会委员名单 .....	(封二)
稿 约 .....	(封三)
作物基因组与数量遗传学专辑(Special issue on crop genome and quantitative genetics)	

综述 (Reviews)

高通量测序技术在多倍体作物基因组学研究中的应用(英文) .....	朱乾浩, 丹尼·卢埃琳, 印·威尔逊(355)
Use of next-generation sequencing in genomic studies of polyploid crops: cotton as an example .....	Zhu Qianhao, Danny Llewellyn, Iain Wilson (355)
植物非编码小 RNA 研究进展(英文) .....	沈恩惠, 刘扬, 叶楚玉, 樊龙江(370)
Recent studies on non-coding small RNAs in plants .....	Shen Enhui, Liu Yang, Ye Chuyu, Fan Longjiang (370)
数量性状位点定位和全基因组关联分析的方法与软件综述(英文) .....	马蒙, 朱军(379)
Tools for quantitative trait locus mapping and genome-wide association study mapping: a review .....	Md. Mamun Monir, Zhu Jun (379)

研究论文 (Research articles)

基于四交群体的连锁图谱构建(英文) .....	蒋蓓蓓, 余世洲, 肖炳光, 楼向阳, 徐海明(387)
Constructing linkage map based on a four-way cross population .....	Jiang Beibei, Yu Shizhou, Xiao Bingguang, Lou Xiangyang, Xu Haiming (387)
东乡野生稻叶绿体基因组拼接及系统进化分析(英文) .....	林张翔, 王莹莹, 付菲, 叶楚玉, 樊龙江(397)
Complete chloroplast genome of Dongxiang wild rice and its application in phylogenetic analysis .....	Lin Zhangxiang, Wang Yingying, Fu Fei, Ye Chuyu, Fan Longjiang (397)
茶树“龙井 43”叶绿体基因组测序及其系统进化(英文) .....	叶晓倩, 赵忠辉, 朱全武, 王莹莹, 林张翔, 叶楚玉, 樊龙江, 须海荣(404)
Entire chloroplast genome sequence of tea ( <i>Camellia sinensis</i> cv. Longjing 43): a molecular phylogenetic analysis .....	Ye Xiaoqian, Zhao Zhonghui, Zhu Quanwu, Wang Yingying, Lin Zhangxiang, Ye Chuyu, Fan Longjiang, Xu Hairong (404)
吸烟对控制 FEV1/FVC 比率的基因上位性及性别互作效应的影响(英文) .....	许昌巍, 朱军(413)
Impacts of cigarette smoking on epistasis and gender-specific effects of FEV1/FVC ratio in human .....	Xu Changwei, Zhu Jun (413)
吸烟和性别对身体质量指数遗传结构的影响(英文) .....	汤斌, 朱军(421)
Impact of cigarette smoking and gender on genetic architecture of body mass index .....	Tang Bin, Zhu Jun (421)
全基因组关联分析 SGRQ 遗传变异与吸烟的关系(英文) .....	郝心颖, 朱军(431)

Genome-wide association studies for identifying genetic architecture of SGRQ in smoking population ..... Hao Xinying, Zhu Jun (431)

烟草核心种质库构建及遗传多样性研究(英文) ..... 周佳萍,杨春元,吴春,王仁刚,史跃伟,谢升东,王志红,徐海明,任学良(440)

Genetic diversity and development of core collection in tobacco (*Nicotiana tabacum* L.) resources ..... Zhou Jiaping, Yang Chunyuan, Wu Chun, Wang Rengang, Shi Yuewei, Xie Shengdong, Wang Zhihong, Xu Haiming, Ren Xueliang (440)

玉米C型细胞质雄性不育系及保持系基因差异表达分析 ..... 王继玥,汪生庆,曹墨菊(451)

Different expression analysis of genes between C type cytoplasmic male sterile line and its maintainer line in maize ..... Wang Jiyue, Wang Shengqing, Cao Moju (451)

水稻 *Ox09g24220* 基因插入突变体的分子鉴定与农艺性状分析 ..... 袁兵,崔海瑞,富昊伟,蒋萌,李瑞清,赵海军,舒庆尧(456)

Molecular characterization and agronomic trait analysis of rice *Ox09g24220* gene insertion mutants ..... Yuan Bing, Cui Hairui, Fu Haowei, Jiang Meng, Li Ruiqing, Zhao Haijun, Shu Qingyao (456)

建兰转录本的微卫星序列和单核苷酸多态性信息分析 ..... 李小白,向林,罗洁,秦德辉,孙崇波(463)

Analysis of microsatellite and single nucleotide polymorphism within transcriptomic database in *Cymbidium ensifolium* ..... Li Xiaobai, Xiang Lin, Luo Jie, Qin Dehui, Sun Chongbo (463)

补 白(Filler)

欢迎订阅《浙江大学学报(农业与生命科学版)》 ..... (378)

新书架 ..... (412)

CN33-1247/S \* 1956 \* b \* A4 \* 120 \* zh+en \* P \* ¥20.00 \* 1000 \* 13 \* 2014-07

责任编辑 胡清华  
英文校对 I. H. Shamsi

## Entire chloroplast genome sequence of tea (*Camellia sinensis* cv. Longjing 43): a molecular phylogenetic analysis

Ye Xiaoqian<sup>1</sup>, Zhao Zhonghui<sup>1</sup>, Zhu Quanwu<sup>1</sup>, Wang Yingying<sup>2</sup>, Lin Zhangxiang<sup>2</sup>, Ye Chuyu<sup>2</sup>, Fan Longjiang<sup>2\*</sup>, Xu Hairong<sup>1\*</sup> (1. Institute of Tea Science, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China; 2. Zhejiang Key Laboratory of Crop Germplasm/Department of Agronomy, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China)

**Summary** *Camellia sinensis* cv. Longjing 43 is a domestic variety of tea species and an important economic crop in China. In this study, we developed a rapid method to get the chloroplast (cp) genome and sequenced the entire cp genome sequence of *C. sinensis* cv. Longjing 43. The *C. sinensis* cv. Longjing 43 cp genome was 157 085 bp in length, which contained a large single-copy (LSC, 86 642 bp) region, a small single-copy (SSC, 18 283 bp) region, and two inverted repeat (IR, each with a size of 26 080 bp) regions. With the cp genome of Korean *C. sinensis* cultivar as a reference, 134 chloroplast genes were successfully annotated. There were 15 genes with non-synonymous mutations in the coding region and more than 100 polymorphic sites in the non-coding region, which could be the DNA markers for the determination of different *C. sinensis* varieties. We also investigated the relationship of 12 *C. sinensis* varieties in China based on several cp genomic regions, which contain many variant sites. The result showed that these varieties were divided into two groups with Lingyunbaimaocha in one group and the other 11 in another group. Among the other 11 varieties, the Longjingchangye, Longjingyuanye, Longjingguazi, and Zhongcha 102 had a closer relationship and were formed into one cluster with 100% support rate, demonstrating the reliability of the method that used the cp genome sequences to investigate the genetic relationships.

**Key words** *Camellia sinensis* cv. Longjing 43; chloroplast genome; high-throughput sequencing; genetic relationship; phylogenetic tree

**CLC number** S 571.1; Q 7 **Document code** A

---

茶树“龙井 43”叶绿体基因组测序及其系统进化(英文). 浙江大学学报(农业与生命科学版), 2014, 40 (4):404-412

叶晓倩<sup>1</sup>, 赵忠辉<sup>1</sup>, 朱全武<sup>1</sup>, 王莹莹<sup>2</sup>, 林张翔<sup>2</sup>, 叶楚玉<sup>2</sup>, 樊龙江<sup>2\*</sup>, 须海荣<sup>1\*</sup> (1. 浙江大学农业与生物技术学院茶叶研究所, 杭州 310058; 2. 浙江大学农业与生物技术学院农学系/浙江省作物种质资源重点实验室, 杭州 310058)

**摘要** 利用高通量测序技术对中国茶树品种“龙井 43”叶绿体基因组的全序列进行测定, 拼接补洞后再利用 Sanger 测序法对序列进行验证, 最终得到“龙井 43”叶绿体基因组全序列 (GenBank 登录号: KF562708.1). 该基因

---

\* **Corresponding author:** Fan Longjiang, Tel: +86-571-88982730, E-mail: fanlj@zju.edu.cn; Xu Hairong, Tel: +86-571-88982215, E-mail: xuhr@zju.edu.cn

**Biography:** Ye Xiaoqian, E-mail: 21216079@zju.edu.cn

**Received:** 2014-05-05; **Accepted:** 2014-05-24; **Published online:** 2014-07-24

URL: <http://www.cnki.net/kcms/doi/10.3785/j.issn.1008-9209.2014.05.051.html>

组大小为 157 085 bp,其中大单拷贝区的长度为 86 642 bp,小单拷贝区的长度为 18 283 bp,反向重复区的长度为 26 080 bp,共注释叶绿体基因 134 个.与韩国茶树品种叶绿体基因组序列比对,在编码区发现 15 个基因发生了非同义突变,在非编码区有 100 多个多态性位点,这些突变可作为茶树品种鉴定的 DNA 标记.同时,选取 12 个中国茶树品种,对其叶绿体基因组上的特异性片段(ycf1, psbA-trnH, psbK-psbI-psbI)进行测序比对并构建系统进化树.结果表明:12 个茶树品种在进化树上分为 2 个亚群,其中“凌云白毛茶”单独形成一支,其他 11 个品种形成一支;而 11 个品种中,“龙井瓜子”“龙井长叶”“龙井圆叶”“中茶 102”形成一支,自展支持率为 100%.表明利用叶绿体基因片段可有效区分茶树品种间的亲缘关系.

**关键词** 龙井 43; 叶绿体基因组; 高通量测序; 亲缘关系; 系统发育树

The chloroplast is an organelle that plays an important role in photosynthesis<sup>[1]</sup>. It also carries out a number of other functions, including the synthesis of starch, fatty acids, pigments, and amino acids<sup>[2]</sup>. It is a semi-autonomous organelle that has a complete set of genetic system and an independent genome<sup>[3]</sup>.

In 1909, C. E. Correns discovered a phenomenon that does not abide by the Mendelian laws of inheritance in *Mirabilis jalapa*. It was the first report to indicate that chloroplast may contain genetic information. The complete chloroplast (cp) genome sequences of *Nicotiana tabacum*<sup>[4]</sup> and *Marchantia polymorpha*<sup>[5]</sup> were first obtained in 1986. In recent years, the cp genomics has developed rapidly with the application of molecular biotechnologies and DNA sequencing techniques. Using the high-throughput sequencing technologies, several cp genome sequences were obtained, such as bamboo<sup>[6]</sup> and duckweed<sup>[7]</sup>. By the end of December 10, 2013, there were 420 complete cp genome sequences deposited in GenBank. The cp genome sequences are useful for the study of plant phylogenetics<sup>[8-10]</sup> and photosynthesis at the molecular level<sup>[11-12]</sup>.

*Camellia sinensis* is an important crop used to make tea<sup>[13]</sup>. In recent years, many *C. sinensis* varieties with special characteristics including albino or etiolated buds have been found in China, such as Zhejianganji baiyecha, Ningboqiannianxue, and Hunanbaojinhuangjinya. The resources of *C. sinensis* are thus much more abundant, and these varieties stimulate the enthusiasm of the farmers to conduct seed selection on *C. sinensis*. Due to the high heterozygosity, the research on the genome of *C. sinensis* is still in the initial stage. The relatively small cp genome will be helpful in investigating the albinism

or etiolation mechanisms in *C. sinensis*. In this study, we obtained the complete cp genome sequences of Longjing 43, which is an important *C. sinensis* variety, based on high-throughput sequencing data. Furthermore, we investigated the relationship among different varieties of *C. sinensis* using cp genomes.

## 1 Materials and Methods

### 1.1 Materials

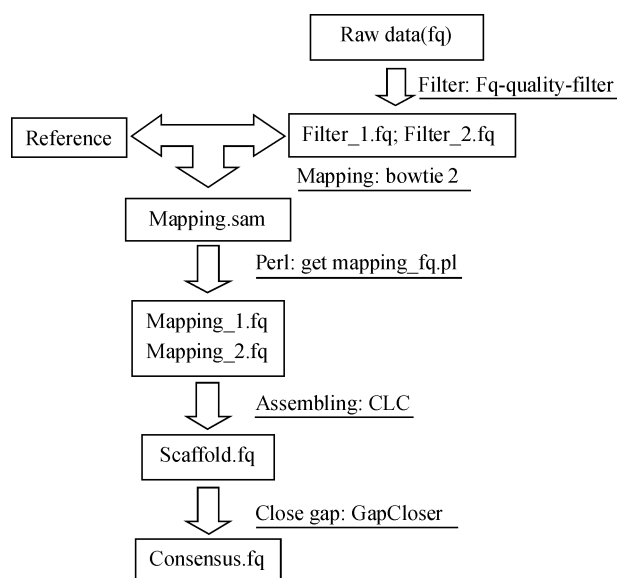
The fresh leaves of Longjing 43, Shuigu tea, Fudingdabai tea, Fujianshuixian, Maoxie, Pingyun, Linyunbaimao tea, Zhenong 25 and Wuniuzao were collected from Huajiachi campus of Zhejiang University. The fresh leaves of Longjingchangye, Longjingguazi, Longjingyuanye and Zhongcha 102 were collected from the Chinese Tea Research Institute.

### 1.2 Methods

**1.2.1 DNA sequencing and chloroplast genome assembly of Longjing 43** The DNA of Longjing 43 was extracted using the plant genome DNA extraction kit (Transgen Company). Raw sequence reads were generated using the Illumina Hiseq 2000 platform and Bowtie 2<sup>[14]</sup> (<http://bowtie-bio.sourceforge.net/index.shtml>) was used to map clean reads with the cp genome of Korean *Camellia* (GenBank number: NC\_020019.1) as a reference. Chloroplast genome-related reads were then collected. Assembly of the mapped reads was then performed using CLC software (<http://www.clcbio.com>). The gaps were filled up by GapCloser<sup>[15]</sup> (<http://soap.genomics.org.cn/soapdenovo.html>) (Fig. 1). Primers used in the PCR amplification were designed by Primer Premier 5.0 (Table 1). PCR products were sequenced using the Sanger method.

**1.2.2 Genome annotation** The cp genome of

Longjing 43 was annotated using DOGMA<sup>[16]</sup> (<http://dogma.ccbb.utexas.edu/>). The final result was submitted to GenBank with the number KF562708.1. The annotated files were used to draw gene maps using the GenomeVx tool<sup>[17]</sup> (<http://wolfe.gen.tcd.ie/GenomeVx/>).



**Fig. 1** Assembly of the Longjing 43 chloroplast genome based on the high-throughput sequencing data

**1.2.3** Phylogenetic relationships among 12 varieties of *C. sinensis* Total DNA was extracted from fresh leaves of the 12 varieties of *C. sinensis* by the CTAB method<sup>[18]</sup>. PCRs for three regions (ycf1, psbA-trnH, psbK-psbI-psbI) with condensed variant sites were carried out in 50  $\mu$ L reaction mixture containing 25  $\mu$ L  $2 \times$  Taq PCR Master Mix, 2  $\mu$ L of each primer (10  $\mu$ mol/L) (Table 1), 19  $\mu$ L ddH<sub>2</sub>O, and 2  $\mu$ L DNA samples. Amplification was carried out under the following condition: 94  $^{\circ}$ C for 4 min, 50 – 55  $^{\circ}$ C (the temperature depends on primers) for 30 s, 72  $^{\circ}$ C for 30 s, recycled around 35 times, and at last 72  $^{\circ}$ C for 4 min. The total DNA samples and PCR products were checked by electrophoresis through 0.8% agarose gel. The PCR products were sent to Shanghai Sonny Biological Technology Co., Ltd to purify and sequence. The raw sequences were checked using Chroms 2.13, and some individual sites were corrected manually via Lasergene. These sequences were then aligned using UltraEdit. The phylogeny reconstruction was

**Table 1** PCR primers used in gap closures and verification of the *C. sinensis* cv. Longjing 43 cp genome

Primer names	Primer locations/bp	Primer sequences (5'—3')
GAP1-F	33 140 – 33 161	AACGCCATGGTAAGGCGTAAGT
GAP1-R	34 116 – 34 095	TCCACGGGTGCAAGAAAGAAAG
GAP2-F	60 382 – 60 404	ATTGCCGAACCCAATGCCTAC
GAP2-R	60 974 – 60 954	TCCCATCCAACAGGGATTCTT
GAP3-F	60 959 – 60 978	TCCCTGTTGGATGGGATTCT
GAP3-R	62 256 – 62 237	TCGTAGGAAGACACGACGAT
GAP4-F	112 473 – 112 494	GAACAAGAGGGATCCACCGAAG
GAP4-R	113 752 – 113 731	GGTTATATTACCGATTTTCGC
GAP5&6-F	113 778 – 113 798	GGACCAAAAACAAGCAAGGGG
GAP5&6-R	114 986 – 114 967	GCTCCACTTCCAGTTCCTGT
GAP7-F	114 898 – 114 920	TGAAAAGCCCATACGACGAAG
GAP7-R	116 377 – 116 356	GATTGGGTTGAAAGCGGCAAT
ycf1-F	127 294 – 127 316	TGTTTCGTTGAGGGAACAGATAC
ycf1-R	128 407 – 128 386	TTCTTCAATATCACGGAACGTC
psbA-trnH-F	156 771 – 156 791	CCGGATCTAAGCGTTGGCTAG
psbA-trnH-R	554 – 533	CGTGCTAACCTTGGTATGGAAG
psbK-psbI-psbI-F	7 802 – 7 821	TATGCTCTGGGACGGAAGGA
psbK-psbI-psbI-R	9 126 – 9 107	TGGCTGAGTGGACTAAAGCG



performed using neighbor-joining (NJ) method<sup>[19]</sup> in Mega 5.0. The confidence of each branch was tested by bootstrap 1 000 times.

2 Results

2.1 Chloroplast genome assembly of Longjing 43

Using the cp genome of Korean *Camellia* as a reference, Longjing 43 cp genome-related reads were collected. After assembly and gap closure, the complete cp genome sequences were obtained.

The 23 G cp genome data for Longjing 43 were obtained with the average sequencing depth of 10. Compared with the reference cp genome sequence, there was 99.98% coverage. After assembly by the CLC Genomics Workbench, the

contig N50 was 26 kb and the longest contig was 41 kb. The complete consensus sequence was 157 085 bp after filling the gaps using GapCloser. We designed six pairs of PCR primers (Table 1, Gap1 – Gap7) for gap validation. The results showed that five gaps were the same as the assembly results. When the Gap7 site was sequenced using the Sanger method, it turned out that the length of gap was 30 bp longer than the assembly result. The final cp genome of *C. sinensis* cv. Longjing 43 has been annotated and deposited in GenBank under the accession number KF562708.1.

2.2 Characteristics of Longjing 43 cp genome

The entire length of the cp genome of Longjing 43 was 157 085 bp (Fig.2), including the large single-copy

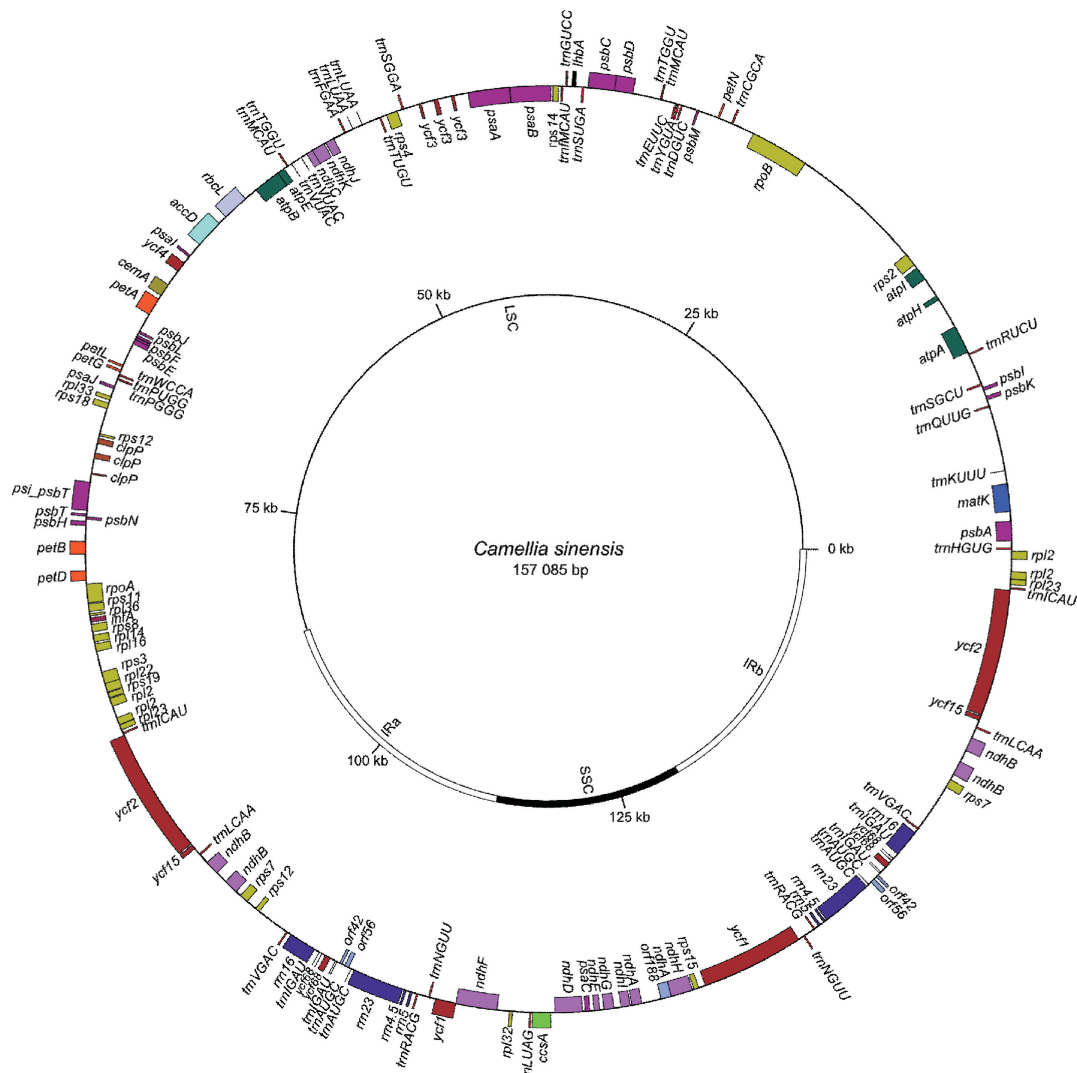


Fig.2 Gene map of the *C. sinensis* cv. Longjing 43 cp genome



(LSC, 86 642 bp) region, the small single-copy (SSC, 18 283 bp) region, and the inverted repeat (IR, 26 080 bp) region. This cp genome consisted of 58% coding region and 42% non-coding region. There were 134 genes divided into 75 self-replication genes, 47 photosynthesis genes, 3 other protein genes and open reading frame, 2 pseudogenes and 7 unknown functional genes (Table 2).

Table 2 Genes of the cp genome of *C. sinensis* cv. Longjing 43

Category for genes	Group of genes	Name of genes
Self replication	tRNA gene	<i>tRNA. UGC</i> * (x2), <i>tRNA. ACG</i> (x2), <i>tRNA. UCU</i> , <i>tRNA. GUU</i> (x2), <i>tRNA. GUC</i> , <i>tRNA. GCA</i> , <i>tRNA. CAU</i> , <i>tRNA. UUG</i> , <i>tRNA. UUC</i> , <i>tRNA. GCC</i> , <i>tRNA. UCC</i> * , <i>tRNA. CAU</i> (x2), <i>tRNA. GUG</i> , <i>tRNA. CAU</i> * (x2), <i>tRNA. CAA</i> (x2), <i>tRNA. UAA</i> * , <i>tRNA. UAG</i> , <i>tRNA. UUU</i> * , <i>tRNA. CAU</i> , <i>tRNA. GAA</i> , <i>tRNA. UGG</i> , <i>tRNA. GCU</i> , <i>tRNA. GGA</i> , <i>tRNA. UGA</i> , <i>tRNA. GGU</i> , <i>tRNA. UGU</i> , <i>tRNA. CCA</i> , <i>tRNA. GUA</i> , <i>tRNA. GAC</i> (x2), <i>tRNA. UAC</i> *
	rRNA genes	<i>rrn4-5</i> (x2), <i>rrn5</i> (x2), <i>rrn16</i> (x2), <i>rrn23</i> (x2)
	Ribosome gene	<i>rpl12</i> * (x2), <i>rpl14</i> , <i>rpl16</i> * , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (x2), <i>rol32</i> , <i>rpl33</i> , <i>rpl36</i> , <i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (x2), <i>rps8</i> , <i>rpsl1</i> , <i>rpsl2-5' end</i> , <i>rpsl2-3' end</i> , <i>rpsl4</i> , <i>rpsl5</i> , <i>rpsl6</i> * , <i>rpsl8</i> , <i>rpsl9</i>
	RNA polymerase gene	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> * , <i>rpoC2</i> *
	Translation initiation factor gene	<i>infA</i>
Genes for photosynthesis	Photosynthesis proteins gene	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i> , <i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
	ATP synthase subunit genes	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> * , <i>atpH</i> , <i>atpI</i>
	NADH subunit gene	<i>ndhA</i> * , <i>ndhB</i> * (x2), <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
	RuBP carboxylase subunit gene	<i>rbcL</i>
	Protease subunit P gene	<i>clpP</i> * *
	Cytochrome-related gene	<i>ccsA</i> , <i>petA</i> , <i>petB</i> * , <i>petD</i> * , <i>petG</i> , <i>petL</i> , <i>petN</i>
Other genes	Membrane protein gene	<i>cemA</i>
	Acetyl-CoA carboxylase subunit gene	<i>accD</i>
	Maturase gene	<i>matK</i>
	Pseudo gene	<i>ψyefl</i> , <i>ψrpsl9</i>
Unknown functional genes		<i>yef1</i> , <i>yef2</i> (x2), <i>yef3</i> * * , <i>yef4</i> , <i>yef5</i> (x2)

\* , \* \* Represent one and two introns, respectively; (x2) indicates the genes duplicated in the cp genome.

2.3 Genetic variation between Longjing 43 and Korean cultivar

The results showed that the length of Longjing 43 cp genome sequence was 18 bp shorter than the Korean cultivar. There were 15 genes with non-synonymous mutation in the coding region ( Table 3 ) of the Longjing 43 cp genome, including gene *rpl2*, which was on the inverted repeat region. Genes *psbT*, *rpl2*, and *rpoC2* of Longjing 43 had frameshift mutations, among which the protein sizes of *psbT* and *rpoC2* were not changed. The mutation of *rpl2* was caused by some insertion, which resulted in adding eight amino

acids to the final protein. There were 16 bases mutated on partial regions of the *ycf1* gene in Longjing 43, which could be a great characteristic among different varieties based on such concentrated single-nucleotide polymorphisms. In comparing the coding regions of the cp genome sequences between Longjing 43 and Korean cultivar, there were more than 100 polymorphic sites found. There especially existed 30 bases inserted in 61 266 sites on the large single-copy of Longjing 43, which could be a DNA marker to identify the different varieties of *C. sinensis*.

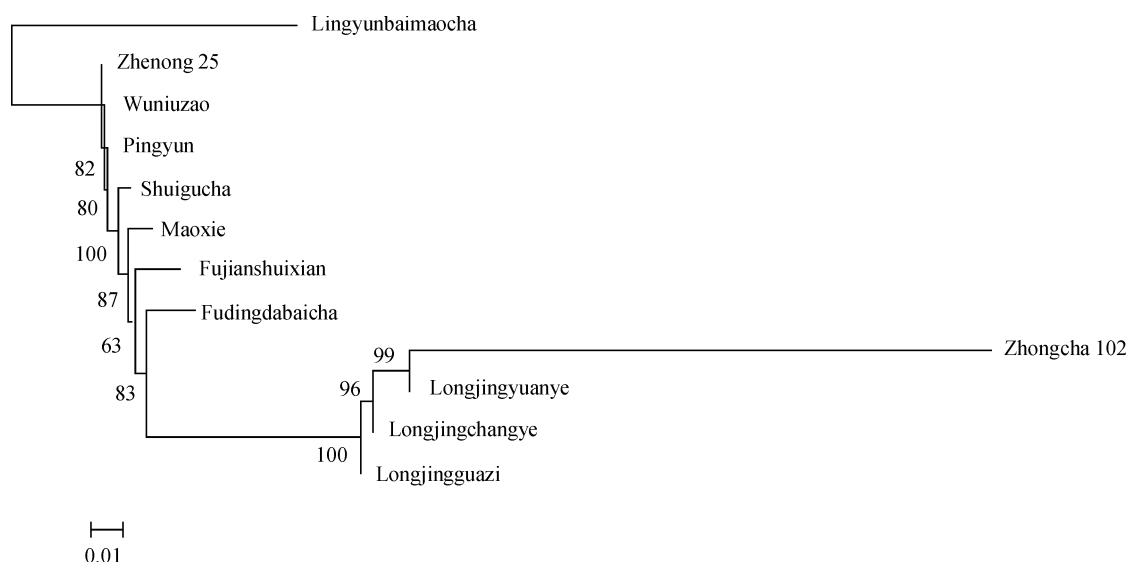
**Table 3 Gene loci with variations between Longjing 43 and Korean cultivar cp genomes**

Gene	Longjing 43 (site; base)	Korean cultivar (site; base)	Differences of Longjing 43 and Korean cultivar
<i>psbT</i>	76 633 – 76 634 76 672 – 76 673	76 641; A 76 680 – 76 681; AG	13AA – 14AA
<i>rpl2</i> (×2)	86 845(IRA); T 156 877(IRB); A	86 848 – 86 849(IRA) 156 896 – 156 897(IRB)	19AA – 24AA
<i>rpoC2</i>	20 084; T 20 165 – 20 166	20 113 – 20 114 20 194; C	28AA – 28AA
<i>matK</i>	2 367; T 3 362; A	2 368; C 3 363; G	N(AAU)– D(GAU) I(AUU)– T(ACU)
<i>rps16</i>	5 205; G	5 208; T	T(ACG)– K(AAG)
<i>rpoC2</i>	18 307; C 20 910; T	18 336; T 20 939; C	A(GCU)– T(ACU) Q(CAA)– R(CGA)
<i>rpoB</i>	26 807; A	26 836; C	V(GUA)– G(GGA)
<i>ycf3</i>	44 549; T	44 587; C	E(GAA)– G(GGA)
<i>accD</i>	59 112; A 59 184; A	59 148; G 59 220; G	K(AAA)– E(GAA) N(AAU)– D(GAU)
<i>clpP</i>	72 526; A	72 532; C	S(UCU)– A(GCU)
<i>rps8</i>	82 856; A	82 866; T	I(AUU)– N(AAU)
<i>ndhD</i>	118 331; G	118 345; T	P(CCA)– T(ACA)
<i>ndhH</i>	124 958; G	124 965; T	Q(CAA)– K(AAA)
<i>rps15</i>	126 018; T	126 025; G	N(AAC)– T(ACC)
<i>ycf1</i>	127 579; G	127 586; A	S(UCG)– L(UUG)
	127 589; G	127 596; A	R(CGU)– C(UGU)
	127 643; T	127 650; G	N(AAU)– H(CAU)
	127 676; T	127 683; A	T(ACG)– S(UCG)
	127 679; G	127 686; A	L(CUU)– F(UUU)
	127 681; T	127 688; A	Y(UAU)– F(UUU)
	127 688; G	127 695; A	L(CUU)– F(UUU)
	127 720; T	127 727; A	Y(UAU)– F(UUU)
	127 754; T	127 761; G	K(AAG)– Q(CAG)
	127 757; T	127 764; G	N(AAC)– H(CAC)
	127 802; G	127 809; A	L(CUU)– F(UUU)
	129 923; A	129 930; G	F(UUU)– L(CUU)

## 2.4 Phylogenetic analysis of 12 varieties of *C. sinensis* in China

The study used the three fragments (ycf1, psbA-trnH, and psbK-psbI-psbI) with condensed variations from 12 varieties of *C. sinensis* to build the phylogenetic tree. The result showed that these varieties were divided

into two groups with Lingyunbaimaocha in one group and the other 11 in another group. Among the other 11 varieties, the Longjingchangye, Longjingyuanye, Longjingguazi, and Zhongcha 102 had a closer relationship and were formed into one cluster with 100% support rate (Fig.3).



The scale indicates the genetic distance.

Fig.3 Phylogenetic tree of 12 tea cultivars based on neighbor-joining method

## 3 Discussion

The second generation sequencing technologies, namely the high-throughput sequencing technologies<sup>[20]</sup>, which preserve the high accuracy of the first generation technologies and can dramatically reduce sequencing costs and increase sequencing speed, serve as a milestone in the field of genomics. A key step in traditional technologies to get cp genome is the preparation of sequencing templates, while the extraction of cpDNA is the limiting factor of the preparation of sequencing templates. Up till now, no general method of extraction was existed. Based on the maternal inheritance and high conservation of cp genome, we collected the cp-related reads using an evolutionarily closed species (Korean *Camellia*) as the reference and then assembled Longjing 43 cp genome. Purification of cpDNA prior to sequencing

is not required in this method. However, if there is a lack of efficient reference sequences or some cp genome fragments are inserted into the nucleic genome, the cp genome is hard to be assembled. In this study, we successfully obtained Longjing 43 cp genome directly from the whole genome sequencing data. It fully demonstrated the efficiency of this method.

Previous results showed that the structure of the cp genome sequences is highly conservative<sup>[21]</sup> and cp genome has important value to the research on the evolution of species and the genetic relationship between different varieties. In study on phylogenetic genomics, we should use the proper DNA sequences with different origins (cp and nucleic genomes), and combine with traditional physiological characters. Then we can obtain the distinctive characters at the different classification levels. This study analyzed the genetic relationship among 12 *C. sinensis* varieties

in China based on the characters of the Longjing 43 cp genome. The results generally matched to the identification theories of species resources<sup>[22]</sup>, which demonstrated the reliability of the method that used the cp genome sequences to investigate the genetic relationships. Recently, DNA molecular markers such as RAPD, AFLD, and RFLP have been used for the determination of *C. sinensis* varieties and the analysis of the genetic polymorphism. Comparing those DNA molecular markers in identifying the relationship among different species, RAPD has a lower reliability, while AFLP and RFLP are more difficult to utilize and expensive. The method presented in this study was based on condensed variant regions of the cp genome, which is easier to be processed and thus has more reliable information.

China was the first country to discover and utilize *C. sinensis*. Korean *C. sinensis* originated from China<sup>[23]</sup>. It can be seen that there are no major differences between the complete cp genome of Longjing 43 and that of Korean *C. sinensis*. There were 15 genes with non-synonymous mutations in the coding region of Longjing 43 cp genome. *psbT* and *rpoC2*, which had been frameshifted, could express proteins in the same size, while *rpl2* expressed the protein with eight more amino-acids. While in the non-coding area, we detected more than 100 polymorphic sites. These mutations may be the result of the genetic changes caused by the environmental changes, which is a possible result of species evolution.

*Camellia sinensis* is an important economic crop in China. Due to the complex chromosome structures, the genomic research on *C. sinensis* is not conducted as much as conducted in other crops. This research took advantages of the high-throughput sequencing technologies to obtain the complete Longjing 43 cp genome, which provided sequence information for further genomic research on *C. sinensis*.

**Acknowledgements** We thank Prof. Chen Liang of Tea Research Institute, Chinese Academy of Agricultural Sciences for offering several materials.

## References:

- [1] Hiratsuka J, Shimada H, Whittier R, *et al.* The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Molecular and General Genetics*, 1989,217:185-194.
- [2] Neuhaus H E, Emes M J. Nonphotosynthetic metabolism in plastids. *Annual Review of Plant Physiology and Plant Molecular Biology*, 2000,51:111-140.
- [3] 王玲,董文攀,周世良.被子植物叶绿体基因组的结构变异研究进展.西北植物学报,2012,32(6):1282-1288.  
Wang L, Dong W P, Zhou S L. Structure mutations and reorganizations in chloroplast genomes of flowering plants. *Acta Botanica Boreali-Occidentalia Sinica*, 2012, 32(6): 1282-1288. (in Chinese with English abstract)
- [4] Shinozaki K, Ohme M, Tanaka M, *et al.* The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *The EMBO Journal*, 1986,5(9):2043-2049.
- [5] Ohyama K, Fukuzawa H, Kohchi T, *et al.* Structure and organization of *Marchantia polymorpha* chloroplast genome. I. Cloning and gene identification. *Journal of Molecular Biology*, 1988,203(2):281-298.
- [6] Zhang Y J, Ma P F, Li D Z. High-throughput of six bamboo chloroplast genomes: phylogenetic implications for temperate woody bamboos (Poaceae: Bambusoideae). *PLoS ONE*, 2011,6(5):e20596.
- [7] Wang W, Messing J. High-throughput sequencing of three Lemnoideae (duckweeds) chloroplast genomes from total DNA. *PLoS ONE*, 2011,6(9):e24670.
- [8] Clarke J L, Daniell H, Nugent J M. Chloroplast biotechnology, genomics and evolution: current status, challenges and future directions. *Plant Molecular Biology*, 2011,76:207-209.
- [9] 姬生国,潘胜利,王峻,等.基于叶绿体 *matK* 基因序列探讨 10 种石杉科石杉属植物的系统关系及分子鉴定.中国中药杂志,2007,32(19):1971-1975.  
Ji S G, Pan S L, Wang J, *et al.* Phylogeny relationship and molecular identification of ten *Huperzia* species (*Huperziaceae*) based on *matK* gene sequences. *China Journal of Chinese Materia Medica*, 2007, 32(19): 1971-1975. (in Chinese with English Abstract)
- [10] Nock C J, Waters D L E, Edwards M A, *et al.* Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal*, 2011,9(3):328-333.
- [11] Whitney S M, Andrews T J. Plastome-encoded bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) supports photosynthesis and growth in tobacco. *Proceeding of the National Academy of Sciences of the United States of America*, 2001,98:14738-14743.
- [12] O'Neill B M, Mikkelsen K L, Gutierrez N M, *et al.* An

- exogenous chloroplast genome for complex sequence manipulation in algae. *Nucleic Acids Research*, 2012, 40(6):2782-2792.
- [13] 马春雷, 陈亮. 茶树功能基因分离克隆研究进展. 分子植物育种, 2006, 4(增刊 3):16-22.
- Ma C L, Chen L. Research progress on isolation and cloning of functional genes in tea plant. *Molecular Plant Breeding*, 2006, 4(Suppl. 3):16-22. (in Chinese with English abstract)
- [14] Langmead B, Trapnell C, Pop M, *et al.* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 2009, 10(3):R25.
- [15] Li R Q, Zhu H M, Ruan J, *et al.* *De novo* assembly of human genomes with massively parallel short read sequencing. *Genome Research*, 2010, 20:265-272.
- [16] Wyman S K, Jansen R K, Boore J L. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*, 2004, 20(17):3252-3255.
- [17] Conant G C, Wolfe K H. GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics*, 2008, 24(6):861-862.
- [18] Tel-Zur N, Abbo S, Myslabodski D, *et al.* Modified CTAB procedure for DNA isolation from epiphytic cacti of the genera *Hylocereus* and *Selenicereus* (Cactaceae). *Plant Molecular Biology Reporter*, 1999, 17:249-254.
- [19] Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 1987, 4:406-425.
- [20] Ansorge W J. Next-generation DNA sequencing techniques. *Nature Biotechnology*, 2009, 25:195-203.
- [21] Bock R. Genetic engineering of the chloroplast: novel tools and new applications. *Current Opinion in Biotechnology*, 2014, 26:7-13.
- [22] 白堃元, 杨亚军, 方嘉禾, 等. 中国茶树品种志. 上海: 上海科学技术出版社, 2001:38-133.
- Bai K Y, Yang Y J, Fang J H, *et al.* *China Tea Varieties*. Shanghai: Shanghai Scientific & Technical Publishers, 2001: 38-133. (in Chinese)
- [23] 金惠淑, 梁月荣, 陆建良. 中、韩两国主要茶树品种基因组 DNA 多态性比较研究. 茶叶科学, 2001, 21(2):103-107.
- Kim H, Liang Y R, Lu J L. Comparative study on genomic DNA diversity between Korean and Chinese tea cultivars by RAPD technique. *Journal of Tea Science*, 2001, 21(2):103-107. (in Chinese with English Abstract)

# 《浙江大学学报(农业与生命科学版)》稿约

1 《浙江大学学报(农业与生命科学版)》是浙江大学主办的有关农业和生命科学类学术期刊(双月刊,国内外公开发行),面向国内外征稿,主要刊登农业科学、生物科学与技术、环境科学、食品科学与技术、核农学与技术、遥测遥感技术及其基础学科和相关学科的学术论文与综述。读者对象为国内外相关的科技工作者、高等院校的教师和研究生等。

## 2 来稿要求及注意事项

2.1 来稿务求论点明确,材料翔实,数据准确可靠,论证合理,结论新颖,文字精炼,语句通顺,可用英文或中文写稿。其中,英文文稿需提供中文题名和摘要等;中文文稿必须包括(按顺序):题名(不超过25字),作者姓名及工作单位(单位须至系、所一级,并包括邮编),中文摘要(内容应包括研究的目的、方法、结果和讨论)及关键词(3~8个,尽量使用名词术语的全称),中图分类号,英文题名,作者姓名的汉语拼音(姓大写,名首字母大写),单位译名,邮编,国别,英文摘要(应写出论文主要试验数据和信息,内容需详细,约500个单词;用第三人称撰写,应符合英语语法)和英文关键词,正文,致谢,参考文献(参考文献务须按本刊格式译成英文和中国学术期刊光盘版的要求著录,外文专著尤其要注意注明出版地和出版者)。此外,在文章首页页脚注明产生论文的基金项目名称和编号,通信作者和第一作者联系电话及E-mail地址。

2.2 来稿请登录本刊网站 <http://www.journals.zju.edu.cn/agr>,点击左侧的“作者投稿”,按提示步骤完成投稿流程(第1次向本刊投稿的作者需先注册,然后投稿)。正文部分用通栏排版,小四号字,务须加宽行距(可采用1.5倍行距),并确保校对无误;采用规范化名词术语,量名称及符号(除雷诺数 $Re$ 等25个特征数和pH值外其他一律用单个字母表示)必须符合国家标准;采用我国法定计量单位(用国际通用符号)。尤其要注意必须将亩换算成公顷( $hm^2$ ),ppm换算为mg/L或mg/kg等。外文字母必须分清大小写和正斜体(注意生物拉丁学名、基因名称和变量用斜体;SI词头、量单位和函数用正体;矩阵、矢量用黑斜体)。

2.3 表格编排采用三线表,不带竖线。图、表要少而精,应有足够的信息,具有自明性,即使使读者不查阅正文即可理解图、表的内容。图、表应排在正文首次提及之后。图题、表题、表中横竖项目栏、注释和图中的内容须中、英文对照。数量关系尽量用图表示,文字部分避免罗列与图、表重复的数字。线条图中点线要准确、清晰、美观,不宜过粗或过细。计算机制图请提供制图参数。显微照片和地图,在图中标出比例尺和指北方向,照片要清晰,反差好。

2.4 参考文献一律采用英汉对照,引文出处按照在文章中出现的先后顺序编号、排列,即按“顺序编码制”编写。在文中引用处的右上角加方括号标明文献序号。文献作者3人以内须全列出,超过3人时列前3人,后加*et al*或等。英美等国作者均按姓前名后顺序书写,名缩写,不加缩写点。作者应对所引文献的准确性负责,修改时请务必逐篇核查文献的正确性。在参考文献中引用非英文文献时,无论其原文是否用英文形式刊出,均用英文形式列出,并在页码后注明语种。中文期刊文献,根据英文摘要的有无在页码后标注(in Chinese with English abstract)或(in Chinese)等。刊名、书名用斜体。几种主要参考文献的著录格式如下:

【期刊】 [序号] 作者.题名.刊名,出版年份,卷(期):起止页码。

【专著】 [序号] 作者.书名.版次(初版不写).译者(原著不写).出版地:出版者,出版年:引文页码。

【论文集】 [序号] 析出文献作者.题名//主编者.论文集名.出版地:出版者,出版年:引文页码。

【学位论文】 [序号] 作者.题名.保存地点:保存单位,年份。

【专利文献】 [序号] 专利所有者.题名:专利国别,专利号.公告日期.获取和访问路径。

【电子文献】 [序号] 主要责任者.题名:其他题名信息[文献类型标志/文献载体标志].出版地:出版者,出版年.获取和访问路径。

\* 期刊:[1] 项延军,李新芝,王小德.5种藤本植物的抗寒性研究初探.浙江大学学报:农业与生命科学版,2011,37(4):421-424.

Xiang Y J, Li X Z, Wang X D. Preliminary study on cold resistance of five different lianas. *Journal of Zhejiang University: Agriculture and Life Sciences*, 2011, 37(4):421-424. (in Chinese with English abstract)

\* 专著:[2] 孙祥生,易现峰.植物生理学实验技术.北京:中国农业出版社,2008:77-79,252-254.

Sun X S, Yi X F. *Plant Physiology Experimental Technique*. Beijing: Chinese Agriculture Publishing House, 2008:77-79, 252-254. (in Chinese)

2.5 来稿如不符上述要求,编辑部均退请作者修改补充后再考虑是否受理。

2.6 来稿经修订,拟于刊用,作者须按标准缴纳版面费。为倡导绿色办刊理念,论文正式刊出后不赠送当期样刊,本刊提供含当期封面、目录的PDF单篇文档,若需纸质样刊需提前订购。为使您的研究成果得到更广泛的传播与交流,您可以给编辑部提供一份名单,我们会把您的全文PDF文件发送给这些专家。

2.7 来稿文责自负。本刊已加入中国科学引文数据库、中国科技论文统计分析数据库、中国学术期刊综合评价数据库、万方数据资源系统数字化期刊群数据库等。本刊发表的所有文章,均同时在上述载体发表。作者需交寄“版权转让协议”。作者著作权使用费与本刊稿酬一次性付给。

《浙江大学学报(农业与生命科学版)》编辑室

封面设计: 万昌平

Zhejiang Daxue Xuebao  
(Nongye yu Shengming Kexue Ban)

本刊被下列重要检索系统收录:

- 美国《化学文摘》(CA)
- 荷兰《文摘与引文数据库》(Scopus)
- 英国《国际农业与生物科学研究中心文摘》(CABI)
- 俄罗斯《文摘杂志》(AJ)
- 波兰《哥白尼索引》(IC)
- 英国《动物学记录》(ZR)
- 英国《食品科技文摘》(FSTA)
- 联合国粮农组织农业索引(AGRIS)
- 美国《乌利希期刊指南》(Ulrichsweb)
- 北京大学图书馆中文核心期刊
- 中国科技核心期刊
- RCCSE 中国核心学术期刊

浙江大学学报  
(农业与生命科学版)  
(双月刊 1956 年创刊)

第40卷 第4期(总第185期) 2014年7月20日

Journal of Zhejiang University  
(Agriculture & Life Sciences)  
(Bimonthly, Started in 1956)

Vol.40 No.4 (General Serial No. 185) July 20, 2014

主管部门: 中华人民共和国教育部	Responsible Unit:	Ministry of Education of the People's Republic of China
主办单位: 浙江大学	Sponsor Unit:	Zhejiang University
主 编: 朱军	Editor in Chief:	Zhu Jun
编辑单位: 《浙江大学学报(农业与生命科学版)》编辑部	Editing Unit:	Editorial Board of Journal of Zhejiang University (Agric. & Life Sci.)
通信地址: 杭州市天目山路148号 邮编310028	Commun. Add:	148 Tianmushan Road, Hangzhou 310028, China
网络地址: <a href="http://www.journals.zju.edu.cn/agr">Http://www.journals.zju.edu.cn/agr</a>	Website:	<a href="http://www.journals.zju.edu.cn/agr">Http://www.journals.zju.edu.cn/agr</a>
电子信箱: <a href="mailto:zdxbsnb@zju.edu.cn">zdxbsnb@zju.edu.cn</a>	E-mail:	<a href="mailto:zdxbsnb@zju.edu.cn">zdxbsnb@zju.edu.cn</a>
电话号码: 0571-88272801	Tel:	86-571-88272801
出版单位: 浙江大学出版社有限责任公司	Publishing Unit:	Zhejiang University Press Co., Ltd.
印刷单位: 杭州丰源印刷有限公司	Printing Unit:	Hangzhou Fengyuan Printing Co., Ltd.
国内发行: 浙江省报刊发行局	Domestic Distribution:	Zhejiang Post Office
订 购 处: 全国各地邮局	Domestic Subscription:	Local Post Office
国外发行: 中国国际图书贸易集团有限公司 (北京市海淀区车公庄西路35号 邮编:100048)	Overseas Distribution:	China International Book Trading Corporation (35 Chegongzhuang Xilu, Haidian District, Beijing 100048, China)