

## The complete chloroplast genome sequence of *Oryza sativa aus*-type variety Nagina-22 (Poaceae)

Yeisoo Yu<sup>a</sup>, Hyun-Oh Lee<sup>a</sup>, Joong Hyoun Chin<sup>b</sup>, Han Yong Park<sup>c</sup> and Soo-Cheul Yoo<sup>d</sup>

<sup>a</sup>Phyzen Genomics Institute, Seongnam, Korea; <sup>b</sup>Graduate School of Integrated Bioindustry, Sejong University, Seoul, Korea; <sup>c</sup>Department of Bioresources Engineering, College of Life Sciences, Sejong University, Seoul, Korea; <sup>d</sup>Department of Plant Life & Environmental Science, Hankyong National University, Ansung, Korea

### ABSTRACT

Rice (*Oryza sativa*) is the predominant staple food crop belonging to the Poaceae family. In this study, complete chloroplast genome sequence of *O. sativa aus*-type variety Nagina-22 was characterized through *de novo* assembly. The genome is a circular DNA molecule of 134,503 bp and has typical quadripartite structures including large single copy region (80,548 bp), small single copy region (12,347 bp), and a pair of inverted repeats (20,804 bp). A total of 120 genes were predicted in the genome, including 77 protein-coding genes, 8 open reading frame genes, 31 tRNA genes, and 4 rRNA genes. Phylogenetic analysis confirmed a close taxonomical relationship with *O. sativa ssp. Indica* species.

### ARTICLE HISTORY

Received 10 November 2017  
Accepted 17 November 2017

### KEYWORDS

*Oryza sativa*; *aus*-type; Nagina-22; N22; chloroplast genome; Poaceae

Rice (*Oryza sativa*) is the most important staple food crop belonging to the Poaceae family. Rice varieties are divided into two major subspecies, *Indica* and *Japonica*, which are further subdivided into five subpopulations, *indica*, *aus*, *temperate japonica*, *tropical japonica*, and *aromatic*. Of these, *indica* and *aus* belong to *Indica* subspecies (Garris et al. 2005; Zhao et al. 2010, 2011; Huang et al. 2012). Rice varieties of *aus* type have been used as valuable donor germplasm for tolerance genes, such as those for submergence-tolerance and phosphorus-deficiency tolerance (Xu et al. 2006; Gamuyao et al. 2012). An *aus*-type rice variety Nagina-22 (or N22) originated from India and shows tolerance against heat and drought (Satake and Yoshida 1978; Prasad et al. 2006; Jagadish et al. 2010; Vikram et al. 2010; Ye et al. 2012; Ye et al. 2015). Therefore, the variety is recognized as a good germplasm for development of new rice cultivar against global warming. In this study, chloroplast genome sequence of rice variety Nagina-22 was completed, which will enrich genetic resources and assist molecular breeding for *aus*-type variety.

Seeds of *O. sativa* cultivar Nagina-22 were originated from International Rice Research Institute (IRRI, GID 4537706 and IRIS 179-1440867) and cultivated by Hankyong National University (IRRI SMTA2017-0072) in Korea.

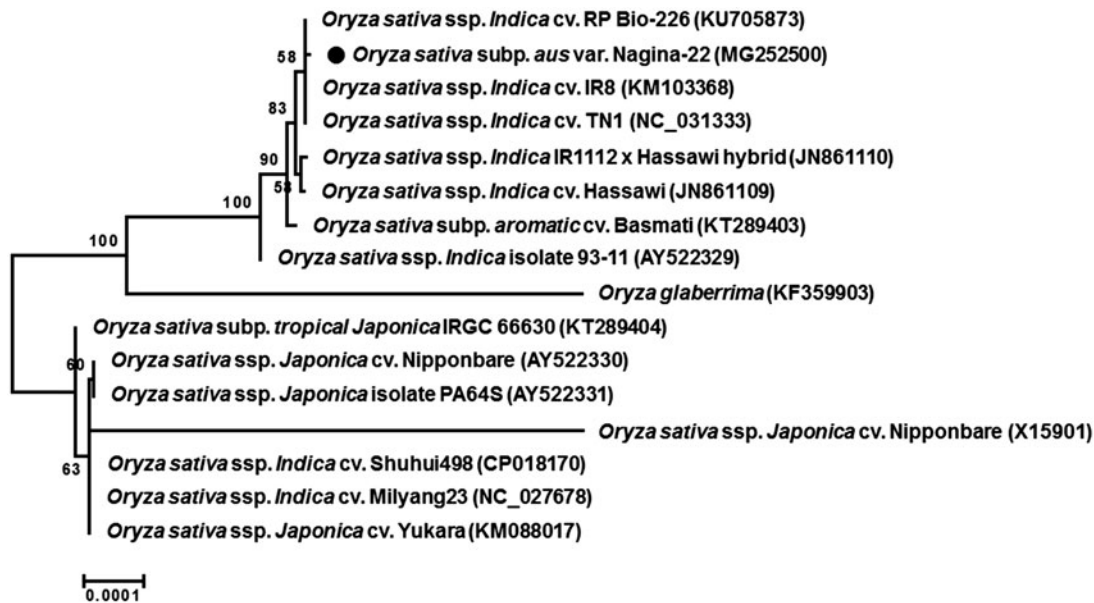
Genomic DNA was isolated from fresh leaves and subjected to constructing an Illumina paired-end (PE) library with 400 bp insert size, according to the manufacturer's instruction. The library was sequenced using an Illumina HiSeq 4000

platform by Macrogen Co. ([www.macrogen.com](http://www.macrogen.com), Seoul, South Korea).

Chloroplast genome sequence was *de novo* assembled using about 977 Mb high quality PE reads and completed using manual correction and gap-filling, as described previously (Kim et al. 2015). The final complete chloroplast genome sequence was annotated using GeSeq (<https://chlorobox.mpimp-golm.mpg.de/geseq-app.html>) and manual curation by Artemis annotation tool (Rutherford et al. 2000) with NCBI BLASTN searches.

Complete chloroplast genome sequence (GenBank Accession no. MG252500) of *O. sativa* variety Nagina-22 is a circular DNA molecule of 134,503 bp, which has typical quadripartite structures consisting of large single copy (LSC) region of 80,548 bp, small single copy (SSC) region of 12,347 bp, and a pair of inverted repeats (IRa and IRb) of 20,804 bp. A total of 120 genes were predicted in the genome, including 77 protein-coding genes, 8 open reading frame genes, 31 tRNA genes, and 4 rRNA genes. Overall GC content of the genome is 39%.

Phylogenetic analysis was performed based on multiple alignments of chloroplast genomes of 16 *Oryza* species. The phylogenetic tree showed that Nagina-22 was grouped much closer with *O. sativa ssp. Indica* species than *ssp. Japonica* species (Figure 1), which was similar to phylogenetic relationship based on nuclear genome SNPs (McNally et al. 2009).



**Figure 1.** Phylogenetic tree of *O. sativa* variety Nagina-22 with related taxa. Whole chloroplast genome sequences were aligned using MAFFT (<http://mafft.cbrc.jp/alignment/server/index.html>) and used to generate maximum-likelihood phylogenetic tree by MEGA 6.0 (Tamura et al. 2013). The numbers in the nodes indicated the bootstrap support values (>50%) from 1000 replicates. Scale bar represents the number of nucleotide substitution per site. GenBank accession numbers of used chloroplast genome sequences are in parentheses. Abbreviation ssp., cv. and subsp. indicate subspecies, cultivar, and subpopulation, respectively.

## Disclosure statement

The authors report no conflict of interest. The authors alone are responsible for the content and writing of the paper.

## Funding

This research was supported by the Next Biogreen 21 program (Project no. PJ01101101) of the Rural Development Administration (RDA), Republic of Korea. This research was also supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2017R1D1A1B04034862).

## References

- Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S. 2012. The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature*. 488:535–539.
- Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics*. 169:1631–1638.
- Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, et al. 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature*. 490:497–501.
- Jagadish SV, Muthurajan R, Oane R, Wheeler TR, Heuer S, Bennett J, Craufurd PQ. 2010. Physiological and proteomic approaches to address heat tolerance during anthesis in rice (*Oryza sativa* L.). *J Exp Bot*. 61:143–156.
- Kim K, Lee SC, Lee J, Yu Y, Yang K, Choi BS, Koh HJ, Waminal NE, Choi HI, Kim NH, et al. 2015. Complete chloroplast and ribosomal sequences for 30 accessions elucidate evolution of *Oryza* AA genome species. *Sci Rep*. 5:15655
- McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, Hoen DR, Bureau TE, et al. 2009. Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc Natl Acad Sci USA*. 106:12273–12278.
- Prasad P, Boote K, Allen L, Sheehy J, Thomas J. 2006. Species, ecotype and cultivar differences in spikelet fertility and harvest index of rice in response to high temperature stress. *Field Crops Res*. 95:398–411.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics*. 16:944–945.
- Satake T, Yoshida S. 1978. High temperature induced sterility in Indica rice at flowering. *Jpn J Crop Sci*. 47:6–17.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 30:2725–2729.
- Vikram P, Singh AK, Singh SP. 2010. Sequence analysis of Nagina-22 drought tolerant ESTs for drought specific SSRs. *Int J Plant Sci*. 5:174–176.
- Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ. 2006. *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature*. 442:705–708.
- Ye C, Argayoso MA, Redoña ED, Sierra SN, Laza MA, Dilla CJ, Mo Y, Thomson MJ, Chin J, Delaviña CB, et al. 2012. Mapping QTL for heat tolerance at flowering stage in rice using SNP markers. *Plant Breeding*. 131:33–41.
- Ye C, Tenorio FA, Argayoso MA, Laza MA, Koh HJ, Redoña ED, Jagadish KS, Gregorio GB. 2015. Identifying and confirming quantitative trait loci associated with heat tolerance at flowering stage in different rice populations. *BMC Genet*. 16:41
- Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, et al. 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat Commun*. 2:467
- Zhao KY, Wright M, Kimball J, Eizenga G, McClung A, Kovach M, Tyagi W, Ali ML, Tung CW, Reynolds A, et al. 2010. Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. *PLoS One*. 5:e10780