Whole genome sequencing data of Oryza sativa L. cv. Takanari

Kajiya-Kanegae, Hiromi ; Yabe, Shiori ; Yoshida, Hiroe ; Ebana, Kaworu ; Yamasaki, Masanori ; Nakagawa, Hiroshi ; Iwata, Hiroyoshi

(URL)
https://hdl.handle.net/20.500.14094/90007256

Data Article

Whole genome sequencing data of *Oryza sativa* L. cv. Takanari

Hiromi Kajiya-Kanegae [a], Shiori Yabe [b], Hiroe Yoshida [c],
Kaworu Ebana [d], Masanori Yamasaki [e], Hiroshi Nakagawa [c],
Hiroyoshi Iwata [a, *]

[a] *Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Science, The University of Tokyo, Japan*
[b] *Institute of Crop Science, NARO, Japan*
[c] *Institute for Agro-Environmental Sciences, NARO, Japan*
[d] *Genetic Resources Center, NARO, Japan*
[e] *Food Resources Education and Research Center, Graduate School of Agricultural Science, Kobe University, Japan*

A R T I C L E I N F O

A B S T R A C T

Oryza sativa L. 'Takanari' is one of the most productive *indica* cultivars [1,2]. Reciprocal chromosome segment substitution lines (CSSLs) derived from a cross between 'Koshihikari' and 'Takanari' are useful tools for the detection and precise mapping of target quantitative trait loci (QTL) in 'Takanari'. Although the available Os-Nipponbare-Reference-IRGSP-1.0 reference genome is available and useful for evaluating genetic diversity among *japonica* cultivars, it is not always useful for evaluating genetic diversity harbored by *indica* cultivars such as 'Takanari'. To reveal sequence variants in 'Takanari' and to exploit these variants in rice breeding programs, the whole genome of 'Takanari' was sequenced using a combination of Illumina HiSeq X Ten (20,983,495 reads and %GC 43) and PacBio (2,847,220 high-quality subreads). NGS data obtained have been deposited in the DNA Data Bank of Japan (DDBJ) under accession number DRA007557.

* Corresponding author. Laboratory of Biometry and Bioinformatics, Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo, 113-8657, Japan.
E-mail address: aiwata@mail.ecc.u-tokyo.ac.jp (H. Iwata).

Specifications Table

Subject area	Biology
More specific subject area	Agricultural and Biological Sciences, Genomics
Type of data	Genome sequence data
How data was acquired	PACBIO_SMRT (Sequel) and ILLUMINA HiSeq X Ten
Data format	Raw sequencing reads
Experimental factors	Whole Genome Sequencing of Oryza sativa L. 'Takanari'
Experimental features	Oryza sativa L. 'Takanari' was sequenced with SMRT and short-read sequencing technologies.
Data source location	Tsukuba, Japan
Data accessibility	The sequence data have been deposited in the DNA Data Bank of Japan (DDBJ) Sequence Read Archive, under submission ID DRA007557. http://trace.ddbj.nig.ac.jp/DRASearch/submission?acc=DRA007557

Value of the data
- Genomic data of high-yielding cultivar 'Takanari' can be useful for identifying genes/quantitative trait loci (QTL) [3].
- The data will help to resolve complex repetitive regions, long indels, and structural variants harbored by Oryza sativa L. [5].
- The data is valuable for those working on crop improvement.

1. Data

The rice cultivar 'Takanari' [1] is a semi-dwarf and high yielding variety developed in Japan [2]. 'Takanari' has a large sink size with a high spikelet number per panicle, high source ability and high carbohydrate translocation capacity [6].

Although short read sequencing has high accuracy and throughput, it is difficult to resolve repeat sequences and identify complex structural variations. Additional genomic techniques, such as long-read sequencing [7], are required for detailed characterization of complex genomic regions in rice genomes [4]. This article reports publicly available whole-genome sequences of 'Takanari' with single molecule, real-time (SMRT) and short-read sequencing technologies. Using PacBio SMRT technology, we produced 25.6 Gb of long-read sequencing raw data with 68 × coverage. Using Illumina paired-end whole-genome shotgun sequencing technology, we generated 6.3 Gb of short-read sequencing data from a 150 bp paired-end library with coverage of 16 × coverage.

The whole-genome sequences reported in this study will be useful for the genetic improvement of rice, by allowing the genetics of the superior agronomic characteristics of 'Takanari' to be investigated.

2. Experimental design, materials, and methods

2.1. Plant materials

In 2017, Oryza sativa L. 'Takanari' plants were cultivated in a growth chamber (SANYO MLR-351H) at the National Agriculture and Food Research Organization in Tsukuba, Ibaraki, Japan.

2.2. DNA extraction and quality control

Genomic DNA was extracted from the leaf blade of Oryza sativa L. 'Takanari' using Tiagen Plant Genomic DNA Kit (TIAGEN, China) following the manufacturer's instructions. Quality and quantity of DNA were checked by agarose gel electrophoresis and Qubit Fluorimeter (Thermo Fisher Scientific Inc., USA).

2.3. Library preparation and sequencing

The extracted genomic DNA was prepared using SMRTbellTM and TruSeq Nano DNA Library preparation methods. For the PacBio Sequel System, SMRTbell TM libraries were prepared following the manufacturer's protocol, Template Preparation and Sequencing Guide (2010−2014) (Pacific Biosciences of California, CA, USA). For HiSeq X Ten, the TruSeq Nano DNA Library was generated using

the Truseq Nano DNA HT Sample Preparation Kit (Illumina USA) following the manufacturer's recommendations.

Acknowledgments

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] T. Imbe, Y. Akama, A. Nakane, T. Hata, K. Ise, I. Ando, H. Uchiyamada, N. Nakagawa, H. Furutachi, N. Horisue, M. Noto, Y. Fujita, K. Kimura, K. Mori, K. Takayanagi, Y. Uehara, S. Ishizaka, M. Nakagahra, T. Yamada, Y. Koga, Development of a multipurpose high-yielding rice variety "Takanari, Bull. Natl. Inst. Crop Sci. 5 (2004) 35−51 ([in Japanese with english summary]), https://www.naro.affrc.go.jp/publicity_report/publication/archive/files/5-3.pdf.

[2] R.D. Taylaran, S. Ozawa, N. Miyamoto, T. Ookawa, T. Motobayashi, T. Hirasawa, Performance of a high-yielding modern rice cultivar Takanari and several old and new cultivars grown with and without chemical fertilizer in a submerged paddy field, Plant Prod. Sci. 12 (2009) 365−380, https://doi.org/10.1626/pps.12.365.

[3] T. Takai, T. Ikka, K. Kondo, Y. Nonoue, Y. Arai-Sanoh, S. Yoshinaga, H. Nakano, M. Yano, M. Kondo, T. Yamamoto, Genetic mechanisms underlying yield potential in the rice high-yielding cultivar Takanari, based on reciprocal chromosome segment substitution lines, BMC Plant Biol. 14 (2014) 295, https://doi.org/10.1186/s12870-014-0295-2.

[4] Y. Kawahara, M. de la Bastide, J.P. Hamilton, H. Kanamori, W.R. McCombie, S. Ouyang, D.C. Schwartz, T. Tanaka, J. Wu, S. Zhou, K.L. Childs, R.M. Davidson, H. Lin, L. Quesada-Ocampo, B. Vaillancourt, H. Sakai, S.S. Lee, J. Kim, H. Numa, T. Itoh, C.R. Buell, T. Matsumoto, Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data, Rice 6 (2013) 4, https://doi.org/10.1186/1939-8433-6-4.

[5] A. Rhoads, K.F. Au, PacBio sequencing and its applications, genomics, Proteomics Bioinforma 13 (2015) 278−289, https://doi.org/10.1016/j.gpb.2015.08.002.

[6] R.D. Taylaran, S. Adachi, T. Ookawa, H. Usuda, T. Hirasawa, Hydraulic conductance as well as nitrogen accumulation plays a role in the higher rate of leaf photosynthesis of the most productive variety of rice in Japan, J. Exp. Bot. 62 (2011) 4067−4077, https://doi.org/10.1093/jxb/err126.

[7] E.L. van Dijk, Y. Jaszczyszyn, D. Naquin, C. Thermes, The Third revolution in sequencing technology, Trends Genet. 34 (2018) 666−681, https://doi.org/10.1016/j.tig.2018.05.008.