Breeding Science 71: 125–133 (2021)
doi: 10.1270/jsbbs.20089

Research Paper

Translation of continuous artificial selection on phenotype into genotype during rice breeding programs

Kenji Fujino*[1)], Yoshihiro Kawahara[2,3)], Kanako O. Koyanagi[4)] and Kenta Shirasawa[5)]

[1)] *Hokkaido Agricultural Research Center, National Agricultural Research Organization*, Sapporo, Hokkaido 062-8555, Japan
[2)] *Institute of Crop Science, National Agricultural Research Organization*, Tsukuba, Ibaraki 305-8518, Japan
[3)] *Advanced Analysis Center, National Agricultural Research Organization*, Tsukuba, Ibaraki 305-8602, Japan
[4)] *Faculty of Information Science and Technology, Hokkaido University*, Sapporo, Hokkaido 060-0814, Japan
[5)] *Kazusa DNA Research Institute*, Kisarazu, Chiba 292-0818, Japan

Understanding genetic diversity among local populations is a primary goal of modern crop breeding programs. Here, we demonstrated the genetic relationships of rice varieties in Hokkaido, Japan, one of the northern limits of rice cultivation around the world. Furthermore, artificial selection during rice breeding programs has been characterized using genome sequences. We utilized 8,565 single nucleotide polymorphisms and insertion/deletion markers distributed across the genome in genotype-by-sequencing for genetic diversity analyses. Phylogenetics, genetic population structure, and principal component analysis showed that a total of 110 varieties were classified into four distinct clusters according to different populations geographically and historically. Furthermore, the genome sequences of 19 rice varieties along with historic representations in Hokkaido, nucleotide diversity and F_{ST} values in each cluster revealed that artificial selection of elite phenotypes focused on chromosomal regions. These results clearly demonstrated the history of the selections on agronomic traits as genome sequences among current rice varieties from Hokkaido.

Key Words: genotype-by-sequencing (GBS), rice, northern limit of rice cultivation, local population, single nucleotide polymorphisms (SNPs), whole genome sequence (WGS).

Introduction

DNA markers could facilitate marker-assisted selection (MAS) in practical crop breeding programs. High-density single nucleotide polymorphisms (SNPs) covering the genome could make MAS in local populations with genetically close relationships available using next-generation sequencing (NGS) systems. With these advantages from recent molecular tools, diverse genetic resources can be utilized to improve traits of current varieties by accelerating crop breeding programs. Genetic bases could provide information for breeding and/or adoption of appropriate varieties for stable crop production around the world.

Asian cultivated rice, *Oryza sativa* L., originated from the tropics (Choi *et al.* 2017, Fuller 2011, Huang *et al.* 2012, Yang *et al.* 2012). Extensive efforts by rice breeding programs have contributed to make rice production possible in various climatic conditions at latitudes ranging between 53°N and 40°S (Fujino *et al.* 2019a, Lu and Chang 1980). Little is known such wide adaptability in rice. Hokkaido (41–45°N latitude) is the northern-most region of Japan and one of the northern limits of rice cultivation around the world. The unique adaptability of rice varieties in Hokkaido may have been established 200–300 years ago (Fujino *et al.* 2019a). Today, rice production in Hokkaido is a significant contributor to agriculture in Japan (Fujino *et al.* 2019c).

Previously, we collected rice landraces from Hokkaido as the Hokkaido Landrace Rice Panel (HLP), which was genetically distinct from varieties from other regions of Japan (Fujino *et al.* 2019a). The genetic population structure of varieties from Hokkaido was different to those from other regions in Japan in phylogenetic analyses using populations across Japan (Fujino *et al.* 2019a, Nagasaki *et al.* 2010, Yamamoto *et al.* 2010, Yonemaru *et al.* 2012). Therefore, current rice breeding programs in Hokkaido have focused on good eating quality for the market in Japan as same as those in other regions (Fujino *et al.* 2019c). Furthermore, genetic bases have shifted four times during rice breeding programs in Hokkaido (Fujino *et al.* 2015, Shinada *et al.* 2014). The shifts successfully achieved the phenotypes to meet current human demands (Fujino *et al.* 2017, 2019c).

Communicated by Mallikarjuna Swamy
Received July 15, 2020. Accepted September 25, 2020.
First Published Online in J-STAGE on February 17, 2021.
*Corresponding author (e-mail: kfujino@affrc.go.jp)

The rice genome of a *japonica* rice variety Nipponbare has been completely sequenced (IRGSP 2005). Furthermore, genome sequences of 3,000 accessions have been registered and accelerated rice functional genomics (Fuentes *et al.* 2019, Wang *et al.* 2018). The genetic base in historical processes since rice domestication may shape rice breeding programs. However, little is known about genes for rice improvements originate in rice breeding programs. NGS technologies can elucidate genome-wide variation patterns and the transmission patterns during rice breeding programs. We may have more concern for genes for agronomic traits.

Although the uniqueness of the adaptability and genetic population structure in rice varieties from Hokkaido has been characterized (Fujino 2003, Fujino and Sekiguchi 2005a, 2005b, 2008, Fujino *et al.* 2013, 2015, 2017, 2019a, 2019b, 2019c, Fujino and Ikegaya 2020, Nonoue *et al.* 2008, Shinada *et al.* 2014), it is unclear whether artificial selection has shaped the genome sequence depending on phenotypes during current rice breeding programs in Hokkaido (Fujino *et al.* 2019c). Here, we demonstrated the genetic relationships of current varieties bred in Hokkaido with Koshihikari, a famous rice cultivar in Japan (Kobayashi *et al.* 2018). Furthermore, the genomes of 19 rice varieties from rice breeding programs in Hokkaido were re-sequenced. We were able to perform genetic characterization of the rice varieties in Hokkaido with genetically close relationships. The results in this study provide understanding of the genomes for rice improvement and provides insights into molecular events in rice breeding programs in the local population in Hokkaido, Japan.

Materials and Methods

Plant materials

To elucidate the characteristics of the current varieties in rice breeding programs, a total of 110 rice varieties in different populations, both geographically and historically, were used for genetic population structure analysis (**Table 1**, **Supplemental Table 1**). We used panels of local varieties, which have already been successfully used to characterize their genetic diversity: the Hokkaido Landraces Panel (HLP) (Fujino *et al.* 2019a) and Hokkaido Rice Core Panel (HRCP) (Fujino *et al.* 2015, 2017, Shinada *et al.* 2014). In addition, 14 varieties bred in 1990–2010 in rice breeding programs in Hokkaido were used as current varieties, CV01–14. As references, eight varieties of ancestral-type rice varieties in Japan, AJ1-01 to AJ1-08, were used. Also, the genome sequence data of eight varieties of KSH including Koshihikari and its progenies (WGS06-14), and Japanese landraces (WGS01–05) in the DRA/DDBJ were used (**Table 1**, **Supplemental Table 1**) (Arai-Kichise *et al.* 2011, 2014).

Furthermore, 19 rice varieties in the HRCP were re-sequenced (**Table 2**). In addition, the genome of a rice variety Cody (JP14658 in Genebank) was sequenced, which may have significant as a parent of exotic germplasm compared with varieties from the rice breeding programs in Hokkaido (Fujino *et al.* 2019d, Shinada *et al.* 2014). As references, the genome sequences of Nipponbare and Kasalath were used.

Table 1. Rice varieties from different populations used in this study

Population		Number of varieties	Dataset
Name	Abbreviation		
Hokkaido Landrace	HL	49	GBS
Hokkaido Rice Core Panel	HRCP	25	GBS
Current Varieties from Hokkaido	CV	14	GBS
Ancestors in Japan 1	AJ1	8	GBS
Koshihikari and its progenies	KSH	9	WGS
Ancestors in Japan 2	AJ2	5	WGS

GBS; genotyping by sequencing, WGS; whole genome sequence.
HRCP was collected in Shinada *et al.* (2014).
WGS in AJ2 is cited from Arai-Kichise *et al.* (2011, 2014).
All varieties in this table are listed in **Supplemental Table 1**.

Seeds of rice varieties were provided by the Genebank of NARO (Tsukuba, Japan) and the Local Independent Administrative Agency Hokkaido Research Organization Hokkaido Central Agricultural Experiment Station (Takikawa, Japan).

Full methods, including DNA analysis, Genotype-by-sequencing (GBS), and Whole-genome sequencing (WGS), are available in the **Supplemental Text**, and these were carried out using standard procedures as described previously. Sequence data from this article have been deposited in the EMBL/GenBank Databases under accession numbers DRA008936, DRA006061, and DRA008447.

Results

Genetic population structure of rice varieties

A total of 290.0 million reads, 29.0 Gb, from 96 rice accessions were sequenced by GBS. The mean was 3.0 million reads per variety. A total of 110 varieties, 96 GBS and 14 WGS, were used for phylogenetic analysis. After filtering, 8,565 polymorphisms including SNPs/insertions/deletions were obtained (**Supplemental Table 2**, **Supplemental Fig. 1**). A dendrogram clearly showed that there were the four clusters, A–D (**Fig. 1**). The four clusters corresponded well with the four populations obtained with $K = 3$ in STRUCTURE analysis (**Fig. 1**). Cluster A involved 11 varieties each from group V in the HRCP and CV. Cluster B was divided into two sub-clusters, B1 and B2. B2 involved seven varieties from Koshihikari and its progenies (KSH), whereas B1 involved 12 varieties from AJ1 and AJ2 and a landrace in Hokkaido, Sakatawase. Most of the HL varieties were divided into two major clusters, C and D. Among 33 varieties in cluster C, 32 were in the HL. Cluster D had 32 varieties, which showed the genome of admixture-type combined HL with rice varieties in the HRCP (**Fig. 1**, **Supplemental Table 1**).

PCA corresponded well with the four clusters in the

Table 2. Genome sequences used in this study

Variety	Origin	Year	Group	Cross combination	Raw		Mapped			Cluster	Reference
					Number of reads	Total reads (Gb)	Number of reads	Total reads (Gb)	Coverage (depth ≥10) (%)		
Hinohikari	Miyazaki	1986	NA	Koganebare/Koshihikari	220,459,074	22.3	215,336,872	20.6	97.4	KSH	This study
Hitomebore	Miyagi	1988	NA	Koshihikari/Hatsuboshi	230,535,534	23.3	225,153,018	21.5	97.3	KSH	This study
Koshihikari	Fukui	1953	NA	Norin No. 22/Norin No. 1	223,469,594	22.6	218,562,790	20.9	97.4	KSH	This study
Nipponbare	Aichi	1961	NA	Yamabiko/Tyushin 110	212,945,738	21.5	209,177,192	19.9	98.9	KSH	This study
Genkitukushi	Fukuoka	2008	NA	Tsukusiroman/Tsukushiwase	240,218,816	24.3	234,814,390	22.4	97.5	KSH	This study
Tenkomori	Toyama	2004	NA	Toyama No. 36/Tokei 1000	228,077,352	23.0	222,836,763	21.3	97.8	KSH	This study
Tentakaku	Toyama	2000	NA	Hanaechizen/Hitomebore	230,871,728	23.3	225,500,806	21.6	97.7	KSH	This study
Hokkaiwase	Hokkaido	Landrace	I	NA	174,234,132	17.4	166,955,908	16.0	84.0	NA	This study
Akage	Hokkaido	Landrace	I	NA	224,213,548	22.4	218,308,372	20.9	85.3	EA	This study
Hayayuki	Hokkaido	1968	II	Shinei/Norin No. 19	148,750,100	14.9	140,780,310	13.5	88.1	EA	This study
Shinei	Hokkaido	1951	II	Tomoenishiki/Norin No. 20	170,673,592	17.1	163,652,903	15.7	86.2	EA	This study
Kyouwa	Hokkaido	1941	IIIa	Rikuu No. 132/Wasefukoku	163,908,840	16.4	155,670,385	14.9	87.9	EA	This study
Sorachi	Hokkaido	1967	IIIa	Kuiku No. 12/Mimasari	165,631,110	16.6	156,952,918	15.0	86.5	MD	This study
Wasefukoku	Hokkaido	1936	IIIa	Nakateaikoku/Bozu No. 6	185,473,688	18.5	178,854,475	17.1	88.4	EA	This study
Hakutyoumochi	Hokkaido	1989	IIIb	Joikumochi No. 381/Onnemochi	148,074,464	14.8	133,999,192	12.7	89.4	MD	This study
Shimahikari	Hokkaido	1981	IIIb	Koshihomare/Sorachi	193,008,254	19.3	186,246,819	17.9	83.0	MD	This study
Yukara	Hokkaido	1962	IIIb	Kanto No. 53/Eiko	177,269,508	17.7	171,655,969	16.5	82.5	MD	This study
Honoka 224	Hokkaido	1990	IV	Toiku No. 214/Kuiku No. 110//Kuiku No. 114	156,454,666	15.6	150,670,656	14.5	86.5	MD	This study
Nourin No. 15	Hokkaido	1940	IV	Ginbozu/Hashiribozu	195,023,616	19.5	188,587,725	18.1	87.2	MD	This study
Daichinohoshi	Hokkaido	2003	V	Kuiku No. 151/Hoshinoyume	157,379,584	15.7	134,127,993	12.8	88.8	DV	Fujino and Ikegaya 2020
Fukkurinko	Hokkaido	2003	V	Kukei 90242B/Hoshinoyume	164,143,152	16.4	154,336,174	14.8	86.6	DV	This study
Hoshinoyume	Hokkaido	1996	V	Akitakomachi/Dohoku No. 48//Kirara 397	153,358,932	15.3	144,499,239	13.7	90.0	DV	Fujino *et al.* 2018
Kirara 397	Hokkaido	1988	V	Toiku No. 214/Dohoku No. 36	180,014,622	18.0	159,758,190	15.3	90.1	DV	This study
Kitaake	Hokkaido	1983	V	Eikei 7361/Dohoku No. 5	151,800,222	13.7	147,583,031	12.6	94.8	DV	Fujino *et al.* 2015
Nanatsuboshi	Hokkaido	2001	V	Hitomebore/Kukei 90242A//Kuiku No. 150	204,076,874	20.4	182,756,005	17.4	88.5	DV	This study
Kitakurin	Hokkaido	2014	V	Fukei No. 187/Kuiku No. 162//Fukkurinko	183,797,896	18.4	174,908,890	16.8	88.0	DV	Fujino *et al.* 2018
Cody	NA	NA	NA	NA	139,565,130	14.1	136,198,559	13.3	94.1	NA	This study
Kasalath	NA	NA	NA	NA	238,066,444	24.0	220,632,717	21.0	88.2	NA	This study

NA; not available.
Groups I–V indicate the genetical population structure in rice varieties from Hokkaido (Shinada *et al.* 2014).
Clusters are defined by SNPs in **Fig. 3** in this study.

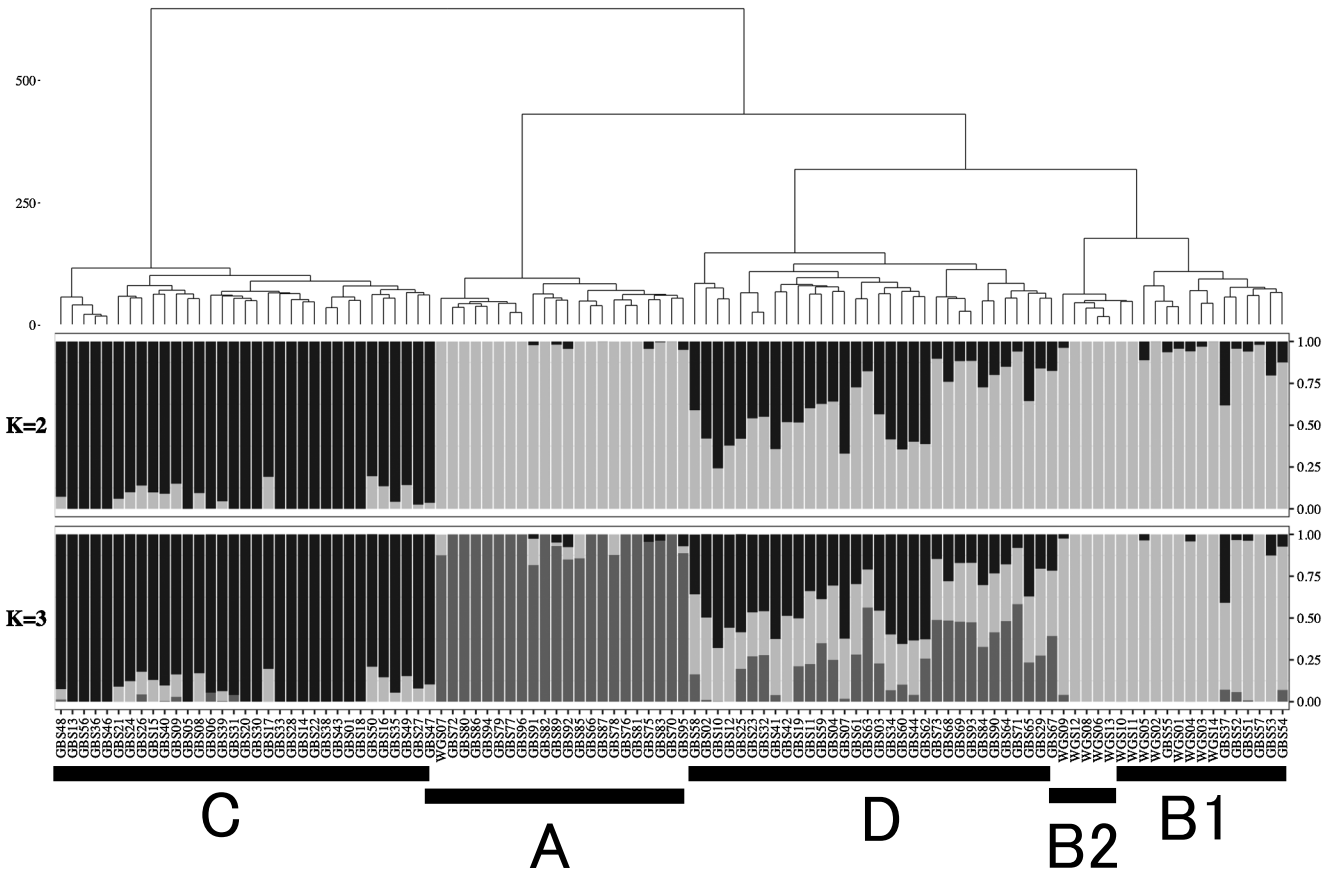

Fig. 1. Classification of 110 rice varieties from different populations with a dendrogram and population structures ($K = 3$) using the 8,565 markers. Bars in the bottom indicate the classifications as clusters A–D.

phylogenetic analysis (**Fig. 2**). The first and second components in PCA explained 18.16% and 10.03% of the total variation, respectively (**Fig. 2**).

Genome sequences of the varieties

Next, we focused on sequence variation over the genome among 29 rice varieties including 19 varieties from Hokkaido. A total of 5.2 billion pair-end reads (526.5 Gb) was sequenced (**Table 2**). Using filtering for read quality, 2,923,374 SNPs were obtained.

A UPGM dendrogram showed that there were four distinct clusters corresponding to historical and geographical differentiation (**Fig. 3**). Clusters early (EA) and middle (MD) involved six and five varieties in the HRCP, respectively. Cluster developed (DV) consisted of seven varieties in the HRCP. Whereas cluster Koshihikari (KSH), including Nipponbare and six varieties of the Koshihikari family, was clearly distinguished from the other clusters. Nipponbare, which is a reference variety in rice research, was grouped into the cluster KSH including Koshihikari. Kasalath, which is an *aus* rice variety, was distinct from the rice varieties in Japan. In addition, Hokkaiwase and Cody were distinct from the rice varieties in Japan.

A heat map was constructed using the number of SNPs and genetic distance among varieties (**Fig. 3A**, **3B**, **Supplemental Table 3**). PCA using 3-d eigenvectors corresponded to the phylogenetic tree (**Fig. 3D**, **Supplemental Table 4**). The genetic groups explained 44.6% of the SNP variation

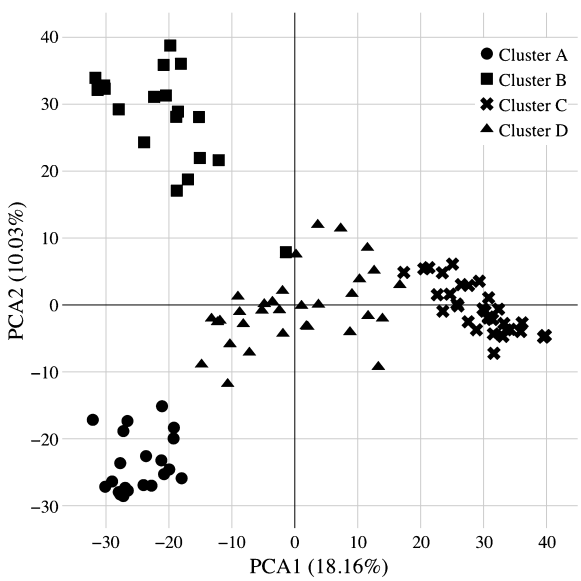

Fig. 2. Principle component analysis (PCA) using genotypes with the 8,565 markers. Clusters A–D correspond to the classification from the phylogenetic analyses in **Fig. 1**.

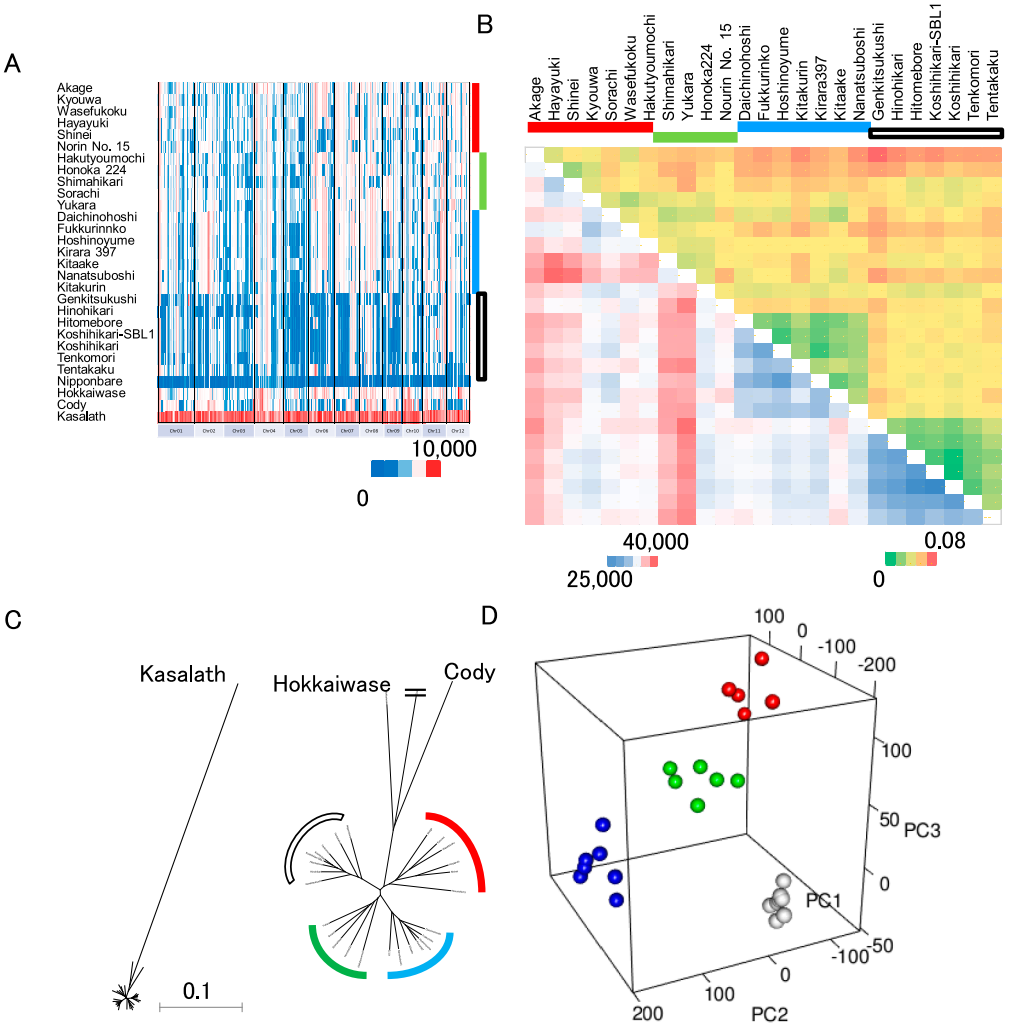

Fig. 3. Classifications of rice varieties by genome sequence variation. SNP density over the genome. The density is expressed for each 1.0 Mb chromosomal region referenced with IRGSP 1.0. (A, B) heat map combined the number of SNPs (lower panel) with distance (upper panel) calculated in Tassel. Values in **Supplemental Table 3** are visualized as color intensity. (C) Phylogenetic tree. Kasalath is an *aus* variety. Hokkaiwase was classified into the cluster of varieties with upland habits (Fujino *et al.* 2019a). Zoom up the branch of the tree for rice varieties from Japan. (D) 3-D plots of PCA. All varieties were classified into four groups corresponding for the three clusters in **Fig. 1**, EA (red), MD (green), and DV (blue) with the reference KSH (white).

in three principal components. The first, second, and third components capturing 19.4%, 16.2%, and 9.1% of variation, respectively.

All phylogenetic approaches concluded that there were three distinct clusters in the Hokkaido rice population along with the historical generations comprising varieties bred between 1936 and 1968 as cluster EA, 1940 and 1981 as cluster MD, and 1983 and 2014 as cluster DV (**Fig. 3D**).

Variations in genome sequences among rice varieties

To evaluate the advances in current varieties in rice breeding programs in Hokkaido, cluster DV, F_{ST} compared between the clusters showed the similarity of genome sequences (**Table 3**). F_{ST} between clusters EA and MD ranged from 0.021 on chromosome 6 to 0.566 on chromosome 12. Whereas, higher F_{ST} was detected between clusters MD and DV. F_{ST} on chromosome was 0.732. Furthermore, 47 1K-SNP-windows showed a high F_{ST} value >0.75. These results suggested that selection on chromosome 2 was intense in the genetic phase change from cluster MD to cluster DV.

SNP distribution

Among 2,923,374 SNPs in the 29 varieties, 2,353,560 (80.5%) were located in intergenic regions and 569,432 (19.5%) were in coding regions (**Supplemental Fig. 2**, **Table 4**). The distributions of SNPs showed cluster specificities; 288 high-impact SNPs were conserved among the three clusters (EA, MD, and DV) in rice breeding programs in Hokkaido. Only 343 SNPs with high impact were shared among all four clusters, which are likely to play an important role in rice cultivation in Japan.

Table 3. Variation in F_{ST} between the clusters over the genome

Chromosome	Number of SNPs	Number of windows	Combination			
			Clusters EA and MD		Clusters MD and DV	
			Average	Number of windows with high F_{ST}	Average	Number of windows with high F_{ST}
1	250,089	250	0.228	1	0.341	5
2	204,069	204	0.133	0	0.732	47
3	213,558	213	0.199	0	0.138	2
4	178,841	178	0.242	0	0.153	0
5	179,616	179	0.186	2	0.156	2
6	185,894	185	0.021	0	0.350	12
7	165,683	165	0.273	11	0.325	14
8	167,268	167	0.224	0	0.215	0
9	133,496	133	0.184	1	0.336	3
10	151,804	151	0.231	0	0.073	0
11	161,669	161	0.200	0	0.208	0
12	139,365	139	0.566	27	0.198	0

High F_{ST} indicates >0.75.
Window shows 1K-SNP-window.

Table 4. Characterization of SNPs

Area	Distribution on the clusters				No. of SNPs			
					Genic			
	EA and MD	DV	KSH	Intergenic	Impact of SNPs			
					High	Moderate	Low	Modifier
1	P	P	P	108,165	343	3,259	3,183	21,722
2	P	P	A	112,184	288	2,761	2,787	21,492
3	A	P	P	10,812	35	295	393	3,618
4	P	A	P	27,737	90	936	917	5,961
5	P	A	A	141,798	329	3,638	3,495	27,447
6	A	A	P	35,471	94	909	1,045	7,677
7	A	P	A	30,360	52	675	666	5,598
8	A	A	A	1,891,353	3,955	38,936	42,782	365,002

AREA is shown as figure in **Supplemental Fig. 2**.
P and A indicate the presence/absence of SNPs in AREAs.

SNPs under selection

To elucidate the role of Cody as exotic germplasm for rice breeding programs in Hokkaido local populations, the genome sequences of seven varieties in cluster DV (**Supplemental Fig. 3**) were compared with that of Cody (**Fig. 4**, **Table 5**). The Kitaake SNPs were present ranging from 31.4% in the Daichinohoshi genome to 51.0% in the Hoshinoyume genome (**Supplemental Fig. 4**, **Table 5**). Among them, 13,927 SNPs in Fukkurinko to 14,673 SNPs in Kirara397 were shared with Cody, which were located on chromosome 2 (**Fig. 4**, **Table 5**).

Discussion

Plant breeding programs have been carried out using selection for desirable phenotypes. Molecular technologies including MAS and NGS may enhance accurate selection based on the genotype for a desirable phenotype. Previously, we identified QTLs/genes responsible for adaptability to Hokkaido (Fujino 2003, Fujino and Sekiguchi 2005a, 2005b, 2008, Fujino *et al.* 2013, 2019a, 2019b, 2019c, Fujino and Ikegaya 2020, Nonoue *et al.* 2008). In addition, we demonstrated phenotypic gains in the 100 years of rice breeding programs in Hokkaido (Fujino *et al.* 2017). Here, we proved the selection of desirable phenotypes has impacted on genome sequences during rice breeding programs in Hokkaido. The gene-based comparison of rice varieties showed they are differentiated both geographically and historically and may elucidate the genome-wide effects of artificial selection.

Genome-wide polymorphisms in this study classified a total of 110 varieties from geographically and historically distinct populations in Japan into four distinct clusters

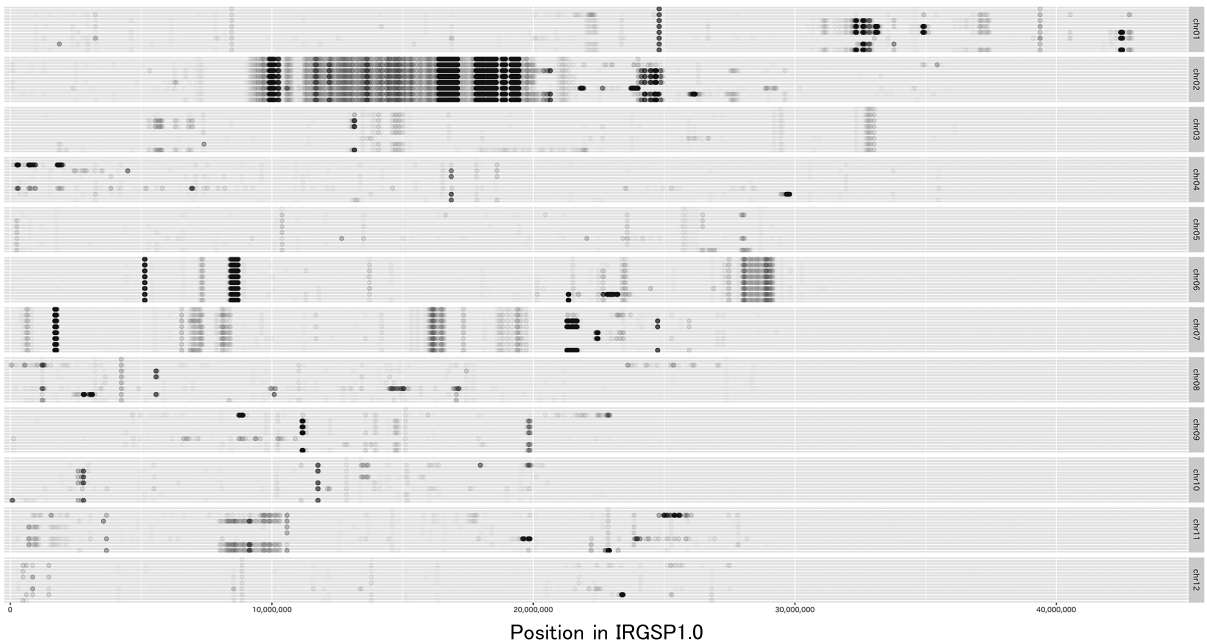

Fig. 4. "Cody" SNPs in varieties of cluster DV. SNPs common to Cody are visualized with a transparency parameter of 0.01, which are highlighted in regions with high-density SNPs. (Top to bottom), All, Daichinohoshi, Fukkurinko, Hoshinoyume, Kirara397, Kitaake, Nanatsuboshi, and Kitakurin. Boxes indicate the chromosomes in rice, chr01–chr12. "All" shows common SNPs among the seven varieties.

Table 5. Impact of Kitaake on the current rice varieties

Variety	No. of SNPs	"Kitaake" SNPs		"Cody" SNPs	
		Number	%	Number	%
Daichinohoshi	72,828	22,846	31.4	14,408	63.1
Fukkurinko	49,577	20,197	40.7	13,927	69
Hoshinoyume	40,911	20,883	51	14,288	68.4
Kitakurin	59,686	20,521	34.4	14,211	69.3
Kirara397	36,056	21,838	60.6	14,673	67.2
Nanatsuboshi	54,817	21,443	39.1	14,087	65.7

No. of SNPs shows SNPs different from those of Nipponbare, Koshihikari and Sorachi. Ambiguous sites were removed.
"Kitaake" SNPs are the same as those of Kitaake but different from those of Cody.
"Cody" SNPs are common in Kitaake and Cody.

(**Figs. 1**, **2**). The genetic population structure may be consistent with historical society in Japan. In the current rice market in Japan, good eating quality is the major human demand, as evidenced by Koshihikari (Fujino *et al.* 2019c, Kobayashi *et al.* 2018). This clustering might be caused by the selection of phenotypes through conventional strategies in rice breeding programs (Fujino *et al.* 2019c). Genetic relationships among these different populations in Japan may depend on potential admixtures, shared ancestry, or pedigrees of local populations (**Figs. 2**, **3**).

Rice breeding programs in Hokkaido were started from landraces in Hokkaido as ancestors in the early 1900s (Fujino *et al.* 2019c). The landraces were classified into two clusters, C and D (**Fig. 1**). Fifteen varieties among the HL population were classified into cluster D, suggesting that they were the founders of rice varieties for rice breeding programs in Hokkaido. They may represent a genetic bottleneck for adaptability to rice cultivation in Hokkaido. Breakthrough of this bottleneck generated in cluster D would reshape rice breeding programs in Hokkaido for human demands in the future (Fujino *et al.* 2019a, 2019c, Shinada *et al.* 2014).

Previously, we demonstrated the significance of improvements in agricultural traits in rice breeding programs in Hokkaido (Fujino *et al.* 2017). The continuous artificial selection on phenotypes during the last 100 years of rice breeding programs in Hokkaido may leave footprints in the genome sequences. Since the early phase of rice breeding programs in Hokkaido, various kinds of traits might have been improved. There may be no signal for intensive selection (**Table 4**), suggesting that these genes might distributed over the genome. Conversely, since the late 1900s rice, good eating quality with stable production could have been placed under strong selection (Fujino *et al.* 2017, 2019c). The genome sequences among the rice varieties in group V of HRCP could conserve a region on chromosome 2 from Cody (**Fig. 4**, **Table 4**). Cody was utilized as a resource for blast disease resistance in rice breeding programs in Hokkaido at that time. Now, molecular studies could show that the resistance gene has been identified as *Pi-cd* on chromosome 11 (Fujino *et al.* 2019d, Shinada *et al.* 2015). The region of chromosome 2 might have been introgressed when Kitaake was selected as new variety. Then, varieties developed from Kitaake carried this conserved region (**Fig. 4**, **Supplemental Fig. 3**), which might establish phenotype in the current varieties.

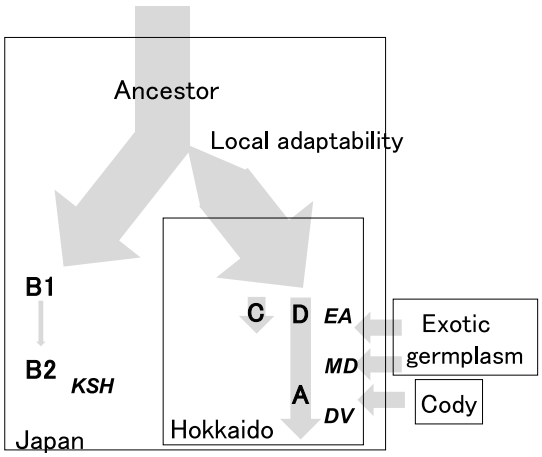

Fig. 5. Model of genetic shifts during adaptation to the northern limits of rice cultivation, Hokkaido. Boxes with Japan/Hokkaido indicate the gene pool of local populations. Bolded A–D show the genetic population structure in the GBS analysis in this study. EA, MD, DV, and KSH, shown in bold and italic, indicate the classification of the genetic population structure in the whole genome sequence analysis in this study. The gray arrow shows the flows of genetic diversity during artificial selections. Mutations for local adaptability to Hokkaido, *ghd7* and *osprr37*, generated rice varieties with extremely early heading date (Fujino *et al.* 2019a).

Combined with our previous work on the unique adaptability of rice at its northern limit of cultivation in Hokkaido, we propose a model for the genetic differentiation of rice populations to meet human demands (**Fig. 5**, **Supplemental Table 5**). Rice cultivation started 2,000–3,000 years ago in Japan. Human communities in Japan have subsequently expanded northwards. Rice with extremely early heading date has been selected for rice cultivation in Hokkaido. Then, rice breeding programs aim to use scientific theory to make rice varieties for human demands as agriculture (Fujino *et al.* 2019c). During this process, tolerance to biotic and abiotic stress has also been improved genetically. Therefore, our study provides new insights and implications for genome-design in practical rice breeding programs.

Author Contribution Statement

Conceived and designed the experiments: KF. Performed the experiments, analyzed the data, wrote the manuscript and approved the final manuscript: KF, YK, KK, and KS.

Acknowledgments

We thank M. Obara (National Agricultural Research Organization) for assistance with DNA experiments, T. Ikegaya (National Agricultural Research Organization) for collecting rice varieties in the HL, and N. Osada (Hokkaido University) for valuable comments. This work was supported in part by a grant from the Iijima Memorial Foundation for the Promotion of Food Science and Technology (to K.F.), the Ministry of Agriculture, Forestry and Fisheries of Japan (Genomics-based Technology for Agricultural Improvement, PFT-1002, to Y.K.), and the Advanced Analysis Center Research Supporting Program of National Agriculture and Food Research Organization.

Literature Cited

Arai-Kichise, Y., Y. Shiwa, H. Nagasaki, K. Ebana, H. Yoshikawa, M. Yano and K. Wakasa (2011) Discovery of genome-wide DNA polymorphisms in a landrace cultivar of *Japonica* rice by whole-genome sequencing. Plant Cell Physiol. 52: 274–282.

Arai-Kichise, Y., Y. Shiwa, K. Ebana, M. Shibata-Hatta, H. Yoshikawa, M. Yano and K. Wakasa (2014) Genome-wide DNA polymorphisms in seven rice cultivars of temperate and tropical *japonica* groups. PLoS ONE 9: e86312.

Choi, J.Y., A.E. Platts, D.Q. Fuller, Y.I. Hsing, R.A. Wing and M.D. Purugganan (2017) The rice paradox: multiple origins but single domestication in Asian rice. Mol. Biol. Evol. 34: 969–979.

Fuentes, R.R., D. Chebotarov, J. Duitama, S. Smith, J.F. De la Hoz, M. Mohiyuddin, R.A. Wing, K.L. McNally, T. Tatarinova, A. Grigoriev *et al.* (2019) Structural variants in 3000 rice genomes. Genome Res. 29: 870–880.

Fujino, K. (2003) Photoperiod sensitivity gene controlling heading date in rice cultivars in the northernmost region of Japan. Euphytica 131: 97–103.

Fujino, K. and H. Sekiguchi (2005a) Mapping of QTLs conferring extremely early heading in rice (*Oryza sativa* L.). Theor. Appl. Genet. 111: 393–398.

Fujino, K. and H. Sekiguchi (2005b) Identification of QTLs conferring genetic variation for heading date among rice varieties at the northern-limit of rice cultivation. Breed. Sci. 55: 141–146.

Fujino, K. and H. Sekiguchi (2008) Mapping of quantitative trait loci controlling heading date among rice cultivars in the northernmost region of Japan. Breed. Sci. 58: 367–373.

Fujino, K., U. Yamanouchi and M. Yano (2013) Roles of the *Hd5* gene controlling heading date for adaptation to the northern limits of rice cultivation. Theor. Appl. Genet. 126: 611–618.

Fujino, K., M. Obara, T. Ikegaya and K. Tamura (2015) Genetic shift in local rice populations during rice breeding programs in the northern limit of rice cultivation in the world. Theor. Appl. Genet. 128: 1739–1746.

Fujino, K., T. Nishimura, H. Kiuchi, Y. Hirayama and T. Sato (2017) Phenotypic changes during 100-year rice breeding programs in Hokkaido. Breed. Sci. 67: 528–534.

Fujino, K., Y. Hirayama, M. Obara and T. Ikegaya (2018) Colocalization of QTLs for hull-cracked rice and grain size in elite rice varieties in Japan. Breed. Sci. 68: 449–454.

Fujino, K., M. Obara and T. Ikegaya (2019a) Establishment of adaptability to the northern-limit of rice production. Mol. Genet. Genomics 294: 729–737.

Fujino, K., U. Yamanouchi, Y. Nonoue, M. Obara and M. Yano (2019b) Switching genetic effects of the flowering time gene *Hd1* in LD conditions by *Ghd7* and *OsPRR37* in rice. Breed. Sci. 69: 127–132.

Fujino, K., Y. Hirayama and R. Kaji (2019c) Marker-assisted selection in rice breeding programs in Hokkaido. Breed. Sci. 69: 383–392.

Fujino, K., Y. Hirayama, M. Obara and T. Ikegaya (2019d) Introgression of the chromosomal region with the *Pi-cd* locus from *Oryza*

meridionalis into *O. sativa* L. during rice domestication. Theor. Appl. Genet. 132: 1981–1990.

Fujino, K. and T. Ikegaya (2020) A novel genotype DATTO5 developed using the five genes exhibits the fastest heading date designed in rice. Breed. Sci. 70: 193–199.

Fuller, D.Q. (2011) Pathways to Asian civilizations: tracing the origins and spread of rice and rice cultures. Rice 4: 78–92.

Huang, X., N. Kurata, X. Wei, Z.X. Wang, A. Wang, Q. Zhao, Y. Zhao, K. Liu, H. Lu, W. Li *et al.* (2012) A map of rice genome variation reveals the origin of cultivated rice. Nature 490: 497–501.

IRGSP (2005) The map-based sequence of the rice genome. Nature 436: 793–800.

Kobayashi, A., K. Hori, T. Yamamoto and M. Yano (2018) Koshihikari: a premium short-grain rice cultivar—its expansion and breeding in Japan. Rice 11: 15.

Lu, J.J. and T.T. Chang (1980) Rice in its temporal and spatial perspectives. *In*: Luh, B.S. (ed.) Rice: Production and Utilization, AVI Publishing Co., Inc., Westport, CT, pp. 1–74.

Nagasaki, H., K. Ebana, T. Shibaya, J. Yonemaru and M. Yano (2010) Core single-nucleotide polymorphisms—a tool for genetic analysis of the Japanese rice population. Breed. Sci. 60: 648–655.

Nonoue, Y., K. Fujino, Y. Hirayama, U. Yamanouchi, S.Y. Lin and M. Yano (2008) Detection of quantitative trait loci controlling extremely early heading in rice. Theor. Appl. Genet. 116: 715–722.

Shinada, H., T. Yamamoto, E. Yamamoto, K. Hori, J. Yonemaru, S. Matsuba and K. Fujino (2014) Historical changes in population structure during rice breeding programs in the northern limits of rice cultivation. Theor. Appl. Genet. 127: 995–1004.

Shinada, H., T. Yamamoto, H. Sato, E. Yamamoto, K. Hori, J. Yonemaru, T. Sato and K. Fujino (2015) Quantitative trait loci for rice blast resistance detected in a local rice breeding population by genome-wide association mapping. Breed. Sci. 65: 388–395.

Wang, W., R. Mauleon, Z. Hu, D. Chebotarov, S. Tai, Z. Wu, M. Li, T. Zheng, R.R. Fuentes, F. Zhang *et al.* (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature 557: 43–49.

Yamamoto, T., H. Nagasaki, J. Yonemaru, K. Ebana, M. Nakajima, T. Shibaya and M. Yano (2010) Fine definition of the pedigree haplotypes of closely related rice cultivars by means of genome-wide discovery of single-nucleotide polymorphisms. BMC Genomics 11: 267.

Yang, J., X. Zhao, K. Cheng, H. Du, Y. Ouyang, J. Chen, S. Qiu, J. Huang, Y. Jiang, L. Jiang *et al.* (2012) A killer-protector system regulates both hybrid sterility and segregation distortion in rice. Science 337: 1336–1340.

Yonemaru, J., T. Yamamoto, K. Ebana, E. Yamamoto, H. Nagasaki, T. Shibaya and M. Yano (2012) Genome-wide haplotype changes produced by artificial selection during modern rice breeding in Japan. PLoS ONE 7: e32982.