Note

DaizuBase, an integrated soybean genome database including BAC-based physical maps

Yuichi Katayose^{*1}), Hiroyuki Kanamori^{1,2}), Michihiko Shimomura³), Hajime Ohyanagi³), Hiroshi Ikawa²), Hiroshi Minami²), Michie Shibata^{1,2}), Tomoko Ito²), Kanako Kurita^{1,2}), Kazue Ito^{1,2}), Yasutaka Tsubokura¹), Akito Kaga¹), Jianzhong Wu¹), Takashi Matsumoto¹), Kyuya Harada¹) and Takuji Sasaki¹)

¹⁾ National Institute of Agrobiological Sciences, 1-2 Ohwashi, Tsukuba, Ibaraki 305-8634, Japan

²⁾ STAFF Institute, 446-1 Ippaizuka, Kamiyokoba, Tsukuba, Ibaraki 305-0854, Japan

³⁾ Mitsubishi Space Software Co. Ltd., Takezono, Tsukuba, Ibaraki 305-0032, Japan

Soybean [*Glycine max* (L) Merrill] is one of the most important leguminous crops and ranks fourth after to rice, wheat and maize in terms of world crop production. Soybean contains abundant protein and oil, which makes it a major source of nutritious food, livestock feed and industrial products. In Japan, soybean is also an important source of traditional staples such as tofu, natto, miso and soy sauce. The soybean genome was determined in 2010. With its enormous size, physical mapping and genome sequencing are the most effective approaches towards understanding the structure and function of the soybean genome. We constructed bacterial artificial chromosome (BAC) libraries from the Japanese soybean cultivar, Enrei. The end-sequences of approximately 100,000 BAC clones were analyzed and used for construction of a BAC-based physical map of the genome. BLAST analysis between Enrei BAC-end sequences and the Williams82 genome was carried out to increase the saturation of the map. This physical map will be used to characterize the genome structure of Japanese soybean cultivars, to develop methods for the isolation of agronomically important genes and to facilitate comparative soybean genome research. The current status of physical mapping of the soybean genome and construction of database are presented.

Key Words: BAC-end sequencing, physical map, database.

Introduction

In 2010, the soybean genome was sequenced and assembled by the Soybean Genome Sequencing Consortium in the USA (Schmutz *et al.* 2010). The genome data are available via databases, phytozome (http://www.phytozome.net/soybean) and Soybase (Grant *et al.* 2010) (http://soybase.org/). Other soybean genomes were sequenced by a next generation sequencer (Kim *et al.* 2010, Lam *et al.* 2010). Soybase is an essential site and tool for soybean researchers to investigate genetics, molecular biology, breeding and genomics. Although this database is important for soybean research, Williams82 genome data are insufficient for Japanese soybean research. We therefore constructed a genome database from the Japanese cultivar Enrei, a common cultivar in Japan. Enrei was selected to construct the physical map and decode the genome sequence.

Communicated by T. Anai

Received October 27, 2011. Accepted December 9, 2011.

*Corresponding author (e-mail: katayose@nias.affrc.go.jp)

BAC library construction

BAC libraries were constructed from nuclear DNA prepared from young leaves of Enrei (Baba *et al.* 2000). Two restriction endonucleases, *Hin*dIII and *Mbo*I, were used for partial digestion of DNA. Partially digested and size-selected DNA (100–180 kb) was ligated into the BAC vector, pIndigoBAC5 (Epicentre Biotechnologies), then transformed into *E. coli*, ElectroMAX DH10B cells (Life Technologies). We picked up 80,000 clones of *Hin*dIII digest, and 100,000 clones of *Mbo*I digest, and designated GMJENa as the *Hin*dIII digest library and GMJENb as the *Mbo*I library. Insert DNAs were 140 and 100 kb for GMJENa and GMJENb libraries, respectively. Each clone was stored in 384-well microplates and kept at -80° C.

End sequencing of BAC clones

Both ends of all clones of GMJENa and 20,000 clones of GMJENb were sequenced by the BigDye Terminator (Life Technologies) method and ABI 3730xl capillary sequencer

662 Katayose, Kanamori, Shimomura, Ohyanagi, Ikawa, Minami, Shibata, Ito, Kurita, Ito, Tsubokura, Kaga, Wu, Matsumoto, Harada and Sasaki

Chromosome	BAC	BAC	Single BAC	Total length	Covered length	Total gap length	Cover rate
	-	contig	contig	(bp)	(bp)	(bp)	
Gm01 (D1a)	4,110	44	6	55,915,595	53,637,206	2,278,389	96
Gm02 (D1b)	3,179	72	10	51,656,713	46,459,754	5,196,959	90
Gm03 (N)	2,462	62	7	47,781,076	43,124,153	4,656,923	90
Gm04 (C1)	2,882	56	6	49,243,852	45,507,841	3,736,011	92
Gm05 (A1)	2,852	39	4	41,936,504	38,979,257	2,957,247	93
Gm06 (C2)	2,760	64	8	50,722,821	45,066,672	5,656,149	89
Gm07 (M)	2,695	48	7	44,683,157	41,367,378	3,315,779	93
Gm08 (A2)	2,763	56	7	46,995,532	43,208,178	3,787,354	92
Gm09 (K)	3,101	40	3	46,843,750	44,090,053	2,753,697	94
Gm10 (O)	3,077	60	6	50,969,635	45,376,931	5,592,704	89
Gm11 (B1)	2,447	49	4	39,172,790	35,810,276	3,362,514	91
Gm12 (H)	2,430	41	5	40,113,140	35,646,507	4,466,633	89
Gm13 (F)	1,992	70	12	44,408,971	36,659,143	7,749,828	83
Gm14 (B2)	3,774	45	6	49,711,204	45,751,866	3,959,338	92
Gm15 (E)	3,117	49	3	50,939,160	47,368,637	3,570,523	93
Gm16 (J)	2,392	49	11	37,397,385	33,708,594	3,688,791	90
Gm17 (D2)	2,363	55	11	41,906,774	37,992,668	3,914,106	91
Gm18 (G)	3,714	63	3	62,308,140	57,128,821	5,179,319	92
Gm19 (L)	2,994	51	5	50,589,441	46,460,750	4,128,691	92
Gm20 (I)	3,893	45	4	46,773,167	43,610,445	3,162,722	93
Total	58,997	1,058	128	950,068,807	866,955,130	83,113,677	91

Table 1. Statistics of "Enrei" BAC-based physical map base on 20 chromosomes

BAC clones mapped on other scaffolds are not shown.

BAC number: number of BAC clones mapped on each chromosome.

BAC contig: number of contigs on each chromosome.

Single BAC contig: number of contigs, consists of one BAC clone.

Total length: base-pair of each chromosome.

Covered length: size of BAC-covered regions.

Total gap length: size of no BAC regions.

Cover rate: (covered length)/(total length) \times 100 (%).

(Life Technologies) (Katagiri *et al.* 2004). The obtained sequence data were analyzed by PhredPhrap software (Ewing and Green 1998, Ewing *et al.* 1998). After exclusion of low-quality (Phred <30) bases, the average read-length of BAC-end sequences was 650 bases.

Mapping of BAC clones and construction of physical map

To identify the physical positions of each sequenced clone, end sequences were analyzed by Blastn with the Williams82 genome assembly (Glyma1.09). After sequencing, endsequenced BAC clones were mapped on each chromosome of the Williams82 genome assembly. Finally, 59361 BAC clones (58997 clones were mapped on 20 chromosomes, 364 clones were mapped on other scaffolds) were mapped on the Williams82 genome and 91% of the genome was covered by Enrei BAC clones (Table 1). We detected differences between Enrei BAC-end sequences and the Williams82 genome assembly. The mismatch rate was 0.2–0.5%, and the deletion rate was less than 0.1% for each chromosome.

DaizuBase

We constructed an integrated soybean genome database, DaizuBase (http://daizu.dna.affrc.go.jp). This database consists of Gbrowse, Unified map and blast search. The Gbrowse page shows BAC-based physical map, unified map page shows linkage map and DNA markers, both are based on Williams82 genome assembly. Gbrowse provides a tracking function for DNA sequence, BAC-end, BAC contigs, GC contents, ESTs, full-length cDNAs (Umezawa *et al.* 2008), DNA markers (Fig. 1). And also, DaizuBase has a sequence, keyword and position search systems.

The prospects

Using the Roche/454 next generation sequencer, GS-FLX Titanium (Margulies *et al.* 2005), 10 equivalent size of the genome of Japanese soybean cultivar, Enrei, has already been sequenced. After analyzing the data, we will upload genome data for Enrei into DaizuBase.

The database will provide SNPs and In/Dels data for Enrei and Williams82 genomes.

Enrei genome data will be useful to distinguish domestic soybean genomes and isolate important genes. Furthermore,

DaizuBase, an integrated soybean genome database

2	TEOD	EPI		Chromosome >> 11213141516171819110	P DEAG1
HOM	E GBrowse	UnifiedM	ap BLAST	Scaffold >> OK	nce of each feature was
DAIZUh	ase		[How to use DAI7] thase]	deduced from the genomic sequence (Giyma Showing 160.6 kibo from C=01	at scattod sequence). nootitions 26 744 849 to 26 905 440
The lat	test update on DAIZ	Ubase (ver.1.1) is now available. *** [Modification history]	Distructions	promone z v, ma, va v z 0,300,440
DAIT	haco is an integrate	l couho-	a conome database and data mining teal consists of 2 ma-	Nevigation: Click one of the rulers to center on a Revenues: Click one of the rulers to center on a Exemples: Carifo Second, Second, Control on a Exemples: Carifo Second, Second, Control on a Exemples: Carifo Second, Second, Control on a Second Second Seco	a location, or click and drag to select a region. Use the Scrol/Zoom bettore to change magnification and position. 2020/2020. Glaw.point.df. Sett. earlieft is '
browser	rs, a gene viewer, ar	nd BLAS	T search system. more	[Bookmark this] [Upload your own dats] [Hide ba	nner] (Share these tracks) (Link to Image) (High-res Image) (Help)
		_	222	CMEDNa0208811 Search Data Source	Paports & Analysis: Download Decessed FASTA File : Configure Co
к	Keyword Search			DA/2Ubase .Cverview	Serel/Zeem: MMN (Show 160,6 kbp) (5) (Fip
Keywo	rd Search			P Bealen	at 201 201 201 201 201 201 201 201 201 201
- enter a k	keyword and choose	from sev	veral parameters to delimit the search.	Details	
Keyword	d:			BIT B Glyma 1 annotation	20196 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198 20198
		Search F	ield ID		Ciparty(156).1 Ciparty(1570.1
Position	n:			D 11 D Glymail TE annotation D 11 D Williams82_BAC	
chromos	some (or scaffold)			D II D BACcord	
Data Set	t:				
Enrei	i-Peking marker	Williams	82 marker BAC_contig BAC_end BAC		
Giym		Giyma1			Provide a second s
search	reset			BIT B FLEDNA	07(4-2)-F0
PC					04/16-36-95 09/16-36-95
				D 13 D Williams82_marker D 13 D EnrolPeking_marker	10
1.0				B II B N region	N.6402.20706342.20706440
				Dear highlighting	
DAI	ZUbase				lbase
DA		UnifiedMa	ip BLAST	D DAIZU	base
DA HOME	GBrowse	UnifiedMa	90 BLAST I lang di magung di Magung ⊡ Mag ⊡ Mag <u>a Vana Kanan Panan Vanan</u>		Browse UnifiedMap BLAST
DA HOME	CUbase GBrowse	UnifiedMa	ap BLAST y lime if were of MC_ent ⊂ MC_entring ⊂ MC <u>-> Vere Enst-Dations Versen</u> Map	D DAIZU HOME C	BBrowse UnifiedMap BLAST
DA HOME	CUbase GBrowse (1 +) Zoom Level (150px/M Linkage Map 2.7501-1179	UnifiedM:	DELAST	D DAIZU HOME C	BBrowse UnifiedMap BLAST Blastn (Querry) Nucleotide (DBI Nucleotide 2) (Williams82 Scaffold (Clyma1) 2)
DA HOME Chromosome (overview	Comparing the second se	UnifiedM:	BLAST 2. Your ExactDates Verse y them of meter of MC_med MC 2. Your ExactDates Verse tree	D DAIZU HOME C Program: DataLib: Your Query Comment:	Blase UnifiedMap BLAST Blastn (Query) Nucleotide (DBI Nucleotide 2 Williams82 Scalfold (Cryma1) 2 2
DAI HOME	Compared and a compar	UnifiedM:	BLAST view direction of Mod_mode Mod_mode And with the second s	D DAIZU HOME C Program: DataLib: Your Query Comment: Your Query:	Browse UnifiedMap BLAST
DAI HOME Chromosome (overview det	Carbon Level 1 (500)/M Linkage May 2.7507 - 14176 3.4007 - 54.322 4.7677 - 54.43 4.7677 - 14176 4.7677 - 14176 4.7777 - 14176 - 14176 - 14176 - 14176 - 14176 - 14176 - 14176 - 14176 - 14176 - 14176 - 14	UnifiedM: D C Displa physical orb	BLAST	D DALZU HOME C Program: Datalbi: Your Query Comment: Your Query:	Browse UnifiedMap BLAST Illiastr (Query) Nucleotide (DB) Nucleotide Williams82 Scalfold (Cymai) 3
DAL HOME Chromosome (overview dot	Carbon Level (1500/M Linkage May 2.7507 - 14170 3.9607 - 54.322 3.9607 - 54.322 3.9607 - 56.322 3.9607	UnifiedM: (b) (c) (c) (c) (c) (c) (c) (c) (c) (c) (c	BLAST view.direction.direct	D DAIZU HOME C Program: DataLib: Your Query Comment: Your Query:	Browse UnifiedMap BLAST
DAI HOME Chromosome (chr(f(SMb)) dvb dvb state state state	Carbon Level (1 (50)) 2.7507 - 14170 2.7507 - 14170 2.7507 - 14170 2.7507 - 14170 2.7507 - 154.021 4.7507 - 54.021 4.7507 - 5002-5010 4.7507 - 5002-5002 4.7507 - 5002-5002 4.75	UnifiedM: (b) (c) Cospia (c) (c) (c) (c) (c) (c) (c) (c) (c) (c)	BLAST A: View EnculPation Viewal View forward with C., with comparison of the forward of the fo	D DAIZU HOME C Program: DataLib: Your Query Comment: Your Query:	Browse UnifiedMap BLAST
DAI HOME Chromosome (overvriew State State State	Carbon Level (1 (50x/M Linkage Mp 2.7587 - 14176 2.7587 - 14176 2.7587 - 14176 2.7587 - 14176 2.7587 - 154.23 4.7587 - 54.23 4.7587 - 154.23 4.7587 - 154.23 4.7597 -	UnifiedM.	BLAST yim of more of MC, and O MC, and O MC 2. View Know/Paking Variant 2:** 1 1:	D DAIZU HOME C Program: DataLib: Your Query Comment: Your Query:	Browse UnifiedMap BLAST Bastn (Guerry Nucleotide DBI Nucleotide Williams82 Scaffed (Clyma) : 2
DAil HOME Chromosome(overvriew 0xe 1986 1987 1988 198	Concernence of the second seco	UnifiedM.	BLAST A-View Constraints Views View Constraints Views Max. Image: A real of the state of the stat	D DALZU HOME C Program: Datalib: Your Query Comment: Your Query:	Browse UnifiedMap BLAST
DAN HOME Chromosome(chr(156Mb) gas state state state state state	Carbon Level (1509x/M Linkage Mo 2,750+	UnifiedM: (0) (1) Oregan (0) (2) Oregan (0) (2) (2) (2) (2) (2) (2) (2) (2) (2) (2	BLAST A-Vest EntriPation Vestor 1 0	D DALZU HOME C Program: DataLib: Your Query Comment: Your Query: or Upload File with FAS (Comer Tig) no file selected	Browse UnifiedMap BLAST
DAN HOME Chromosome(chr(1556b) gas gas <td>Carbon Levie (15000/M Circles 44.43 2.556-1-16170 3.556-1-16170 3.556-1-16170 3.556-1-16170 3.556-1-16170 3.556-1-1000-10007 3.556-1-1000-1000-10007 3.556-1-1000-1000-1000-1000-1000-1000-1000</td> <td>UnifiedM:</td> <td>BLAST </td> <td>D DALZU HOME C Program: DataLib: Your Query Comment: Your Query: Counce Tiple with FAS Counce Tiple on the selected</td> <td>Browse UnifiedMap BLAST</td>	Carbon Levie (15000/M Circles 44.43 2.556-1-16170 3.556-1-16170 3.556-1-16170 3.556-1-16170 3.556-1-16170 3.556-1-1000-10007 3.556-1-1000-1000-10007 3.556-1-1000-1000-1000-1000-1000-1000-1000	UnifiedM:	BLAST	D DALZU HOME C Program: DataLib: Your Query Comment: Your Query: Counce Tiple with FAS Counce Tiple on the selected	Browse UnifiedMap BLAST
DAI HOME Chromosome (Chromosome (Overriew 000 <td< td=""><td>Carbon Level 1 (500x/M Linkage Map 2,7507 - 14170 2,7507 - 14170 2,7507 - 14170 2,7507 - 14170 2,7507 - 14170 2,7507 - 1507 2,7507 - 1507 2,7507 - 1507 2,7507 - 1507 2,7507 - 1507 2,7507 - 1507 2,5507 - 1507 2,5507 2,5507 - 1507 2,5507</td><td>UnifiedM.</td><td>BLAST View ExactPation Views ************************************</td><td>D DALZU HOME C Program: DataLib: Your Query Comment: Your Query: or Upload File with FAS Comer File no file selected <u>Search Rest</u></td><td>Browse UnifiedMap BLAST</td></td<>	Carbon Level 1 (500x/M Linkage Map 2,7507 - 14170 2,7507 - 14170 2,7507 - 14170 2,7507 - 14170 2,7507 - 14170 2,7507 - 1507 2,7507 - 1507 2,7507 - 1507 2,7507 - 1507 2,7507 - 1507 2,7507 - 1507 2,5507 - 1507 2,5507 2,5507 - 1507 2,5507	UnifiedM.	BLAST View ExactPation Views ************************************	D DALZU HOME C Program: DataLib: Your Query Comment: Your Query: or Upload File with FAS Comer File no file selected <u>Search Rest</u>	Browse UnifiedMap BLAST
DAI HOME Chromosome (Chromosome (State 1840 1840 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000	Carbon Control (1997)	UnifiedM.	BLAST View Enuclidation Views DAC.unity	D DALZU HOME C Program: DataLb: Your Query Comment: Your Query: or Upload File with FAS Comment File with FAS Comment File with FAS Comment File with FAS	Browse UnifiedMap BLAST
DAI HOME Chromosome (chromosome (chromosome (state	Construction of the second sec	UnifiedM.	BLAST AVec Excel/basis Vector 1 0<	D DALZU HOME C Program: Datalb: Your Query Comment: Your Query: or Upload File with FAS (Comer File) no file selected Seath Fast Parameters: Expect:	Browse UnifiedMap BLAST Match: 12
DAI HOME Chromosome (Concernence	UnifiedM.	P Aller 1 0.000,000,000,000,000,000,000,000,000,0	D DALLU HOME C Program: DataLib: Your Query Comment: Your Query: or Upload File with FAS or Upload File with FAS Commer file role is selected Genering role in selected Genering role is selected Generi	Browse UnifiedMap BLAST
DAN HOME Chromosome (chr(Stder) state	Construction Construction 1 2000 Low Set 2,2507 - 161,252 2,2507 - 161,252 4,2507 - 161,252 4,2507 - 161,252 4,2507 - 161,252 4,2507 - 161,252 4,2507 - 162,252 4,2507 - 162,252,253 4,507 - 166,253,250 4,507 - 166,253,250 4,507 - 166,254,250 4,508 - 166,254,250 4,508 - 166,254,250 4,508 - 166,254,250 4,508 - 166,254,250 1,258 - 166,257,250 1,258 - 169,257,158 1,258 - 169,257,158 1,258 - 169,257,158 1,258 - 169,257,158 1,258 - 169,257,158 1,258 - 169,257,158 1,258 - 169,257,158 1,258 - 169,257,158	UnifiedM.	PLAST	D DALLU HOME C Program: Datalb: Your Query Comment: Your Query: Cource Tiel on the selected Search Search Parameters: Expect: WordSize: Filter: C Gi:	STA Format Queries:
DAA HOME Chromessme State	CLOBE - 0000-00000 - 000000 - 0000-00000 - 0000000000	UnifiedM	P BLAST View ExactParket View View ExactParket View	D DALLU HOME C Program: DataLib: Your Query Comment: Your Query: Course File with FAS Course File no file selected Search Beart Parameters: Expect: Filter: Gi: Gi: Gi:	Image: State of Control St
DAAl HOME Chromosome (Soverview Soverv	Construction of the second sec	UnifiedM Uni	BLAST View description of Mod_ment (Mod_ment (M	D DALLU HOME C Program: Datalib: Your Query Comment: Your Query Comment: Your Query: or Upload File with FAS Comme File no file selected Earch East Parameters: Expect: Filter: Gl: Gap Align: Descriptions:	Street UnifiedMap BLAST Image: Statistic (Cyma) : 1 : Image: Statistic (Cyma) : 1 : STA Format Queries: : 12 Match: 12 554 Matk : 12 Gap Extend: 12 12 Coport: 12 12 Cap Extend: 12 12 Coport: 12 13 14
	Carbon Level (1997) 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	UnifiedMon poyseal 376	BLAST View Contribution V	D DALLU HOME C Program: Datalb: Your Query Comment: Your Query Comment: Your Query: Comer Table Not FAS Comer Table Not FAS Co	Browse UnifiedMap BLAST
DAAl HOME Chromosome (000 000 1000 <td>Bits Bits <th< td=""><td>UnifiedMode Physical 376</td><td>BLAST View Constraints 1</td><td>D DALLU HOME C Program: Datalbi: Your Query Comment: Your Query: C C Upload File with FAS C Choese File on the selected C C C C C C C C C C C C C C C C C C C</td><td>Browse UnifiedMap BLAST</td></th<></td>	Bits Bits <th< td=""><td>UnifiedMode Physical 376</td><td>BLAST View Constraints 1</td><td>D DALLU HOME C Program: Datalbi: Your Query Comment: Your Query: C C Upload File with FAS C Choese File on the selected C C C C C C C C C C C C C C C C C C C</td><td>Browse UnifiedMap BLAST</td></th<>	UnifiedMode Physical 376	BLAST View Constraints 1	D DALLU HOME C Program: Datalbi: Your Query Comment: Your Query: C C Upload File with FAS C Choese File on the selected C C C C C C C C C C C C C C C C C C C	Browse UnifiedMap BLAST
DAA HOME Lhomesome (Sole	Control Desc. Control Control <thcontrol< th=""> <thcontrol< th=""> <thcon< td=""><td>UnifiedMon) To average of the second secon</td><td>************************************</td><td>D DALLU HOME C Program: DataLib: Your Query Comment: Your Query: C C Upload File with FAS C Descriptions file selected Gearch Test Parameters: Expect: WordSize: Filter: Gap Align: Descriptions: Alignments:</td><td>Browse UnifiedMap BLAST Itastn (Guerri Nucleotide (DBI Nucleotide (Willams2 Staffold (Cyma)) = = STA Format Queries:</td></thcon<></thcontrol<></thcontrol<>	UnifiedMon) To average of the second secon	************************************	D DALLU HOME C Program: DataLib: Your Query Comment: Your Query: C C Upload File with FAS C Descriptions file selected Gearch Test Parameters: Expect: WordSize: Filter: Gap Align: Descriptions: Alignments:	Browse UnifiedMap BLAST Itastn (Guerri Nucleotide (DBI Nucleotide (Willams2 Staffold (Cyma)) = = STA Format Queries:
	Bits Control Control <thcontrol< th=""> <thcontrol< th=""> <thcont< td=""><td>UnifiedM</td><td>P DUART View Construction of Mode on Construction of Construction</td><td>D DALLU HOME C Program: Datalb: Your Query Comment: Your Query Comment: Your Query: or Upload File with FAS Comme File no file selected East Parameters: Expect: Filter: Gl: Gap Align: Descriptions: Alignments:</td><td>Brows UnifiedMa BLAST</td></thcont<></thcontrol<></thcontrol<>	UnifiedM	P DUART View Construction of Mode on Construction of Construction	D DALLU HOME C Program: Datalb: Your Query Comment: Your Query Comment: Your Query: or Upload File with FAS Comme File no file selected East Parameters: Expect: Filter: Gl: Gap Align: Descriptions: Alignments:	Brows UnifiedMa BLAST
	Bits Dist Dist <thdis< th=""> <thdist< th=""> Dist Di</thdist<></thdis<>	Unified Museum 2017 2017 2017 2017 2017 2017 2017 2017	BLAST View Contribution Markar of Dirac of	D DALLU HOME C Program: Balib: Your Query Comment: Your Query Comment: Your Query: Comer Tag) no file selected Search Reset Parameters: Expect: WordSize: Filter: Gl: Gl: Gl: Alignments:	Brows UnifedMa BLAST
DAAl HOME Chromosome (Some 1900 </td <td>Carbon Levie T (1992) 1 3 20m Levie T (1992) 1 4 70m - 4 4.13 1 4 70m - 4 4.13 1</td> <td>UnifiedMum))) () () () () () () () () () () () (</td> <td>BLAST Image: Description of the description of th</td> <td>D DALLU HOME C Program: DataLib: Your Query Comment: Your Query Comment: Your Query: C C Upload File with FAS C Consectific) no file selected C C C C C C C C C C C C C C C C C C C</td> <td>Brows UnifiedMag BLAST Image: Comparison of the comparison of</td>	Carbon Levie T (1992) 1 3 20m Levie T (1992) 1 4 70m - 4 4.13 1	UnifiedMum))) () () () () () () () () () () () (BLAST Image: Description of the description of th	D DALLU HOME C Program: DataLib: Your Query Comment: Your Query Comment: Your Query: C C Upload File with FAS C Consectific) no file selected C C C C C C C C C C C C C C C C C C C	Brows UnifiedMag BLAST Image: Comparison of the comparison of
DAA HOME Chromosome (Some 1 2000 1 2000 1 2000 1 2000	Bits Bits <th< td=""><td>UnifiedMon) To use of the second se</td><td>************************************</td><td>D</td><td>Brows UnifedMax BLAST Image: Comparison of the state of the state</td></th<>	UnifiedMon) To use of the second se	************************************	D	Brows UnifedMax BLAST Image: Comparison of the state
	Bits Bits <th< td=""><td>UnifiedMut) Couperson) Coup</td><td></td><td>D DALLU HOME C Program: Datalb: Your Query Comment: Your Query Comment: Your Query: or Upload File with FAS Commer File no file selected East Parameters: Expect: Filter: Gl: Gap Align: Descriptions: Alignments:</td><td>Stress UnifiedMag BLAST Image: Control Nucleotide (DB) Nucleotide Image: Control Nucleotide Image: Control Nucleotide Image: Control Nucleotide Image: Control Nucleotide STA Format Queries: Image: Control Nucleotide I</td></th<>	UnifiedMut) Couperson) Coup		D DALLU HOME C Program: Datalb: Your Query Comment: Your Query Comment: Your Query: or Upload File with FAS Commer File no file selected East Parameters: Expect: Filter: Gl: Gap Align: Descriptions: Alignments:	Stress UnifiedMag BLAST Image: Control Nucleotide (DB) Nucleotide Image: Control Nucleotide Image: Control Nucleotide Image: Control Nucleotide Image: Control Nucleotide STA Format Queries: Image: Control Nucleotide I

Fig. 1. Browsing DaizuBase. A) DaizuBase top page with links to Gbrowse, Unified Map and Blast search. B) Gbrowse shows BAC-based physical map data. C) Unified Map shows relationships among the linkage map, DNA markers and BAC end sequences. D) Sequence search systems using BLAST.

sequencing of various Japanese cultivar genomes is progressing using the next generation sequencer. These genomic data will be useful for establishing DNA markers for Japanese cultivars.

Acknowledgements

We thank Dr. Naoki Katsura, President of the STAFF Institute, for encouragement and continuous support of the Soybean Genome Research. This work was supported by a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Genomics for Agricultural Innovation, DD-1020 and SOY1001).

Literature Cited

- Baba, T., S. Katagiri, H. Tanoue, R. Tanaka, Y. Chiden, S. Saji, M. Hamada, M. Nakashima, M. Okamoto, M. Hayashi *et al.* (2000) Construction and characterization of Rice genomic libraries: PAC library of Japonica variety, Nipponbare and BAC library of Indica variety, Kasalath. Bulletin of NIAR 14: 41–49.
- Ewing, B. and P.Green (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8: 186– 194.
- Ewing, B., L.Hiller, M.C. Wendel and P.Green (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8: 175–185.

663

664 Katayose, Kanamori, Shimomura, Ohyanagi, Ikawa, Minami, Shibata, Ito, Kurita, Ito, Tsubokura, Kaga, Wu, Matsumoto, Harada and Sasaki

- Grant, D., R.T.Nelson, S.B.Cannon and R.C.Shoemaker (2010) Soy-Base, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res. 38: D843–D846.
- Katagiri, S., J. Wu, Y. Ito, W. Karasawa, M. Shibata, H. Kanamori, Y. Katayose, N. Namiki, T. Matsumoto and T. Sasaki (2004) End sequencing and chromosomal *in silico* mapping of BAC clones derived from an *indica* rice cultivar, Kasalath. Breed. Sci. 54: 273– 279.
- Kim, M.Y., S.Lee, K.Van, T.-H.Kim, S.-C.Jeong, I.-Y.Choi, D.-S. Kim, Y.-S.Lee, D.Park, J.Ma *et al.* (2010) Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. Proc. Natl. Acad. Sci. USA 107: 22032–22037.
- Lam, H.-M., X.Xu, X.Liu, W.Chen, G.Yang, F.-L.Wong, M.-W.Li, W.He, N.Qin, B.Wang *et al.* (2010) Resequencing of 31 wild and

cultivated soybean genomes identifies patterns of genetic diversity and selection. Nature Genet. 42: 1053–1061.

- Margulies, M., M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, J.Berka, M.S.Braverman, Y.J.Chen, Z.Chen *et al.* (2005) Genome sequencing in microfabricated high-density picoliter reactors. Nature 437: 376–380.
- Schmutz, J., S.B. Cannon, J.Schlueter, J.Ma, T.Mitros, W.Nelson, D.L. Hyten, Q.Song, J.J.Thelen, J.Cheng *et al.* (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183.
- Umezawa, T., T. Sakurai, Y. Totoki, A. Toyoda, M. Seki, A. Ishiwata, K. Akiyama, A. Kurotani, T. Yoshida, K. Mochida *et al.* (2008) Sequencing and analysis of approximately 40 000 soybean cDNA clones from a full-length-enriched cDNA library. DNA Res. 15: 333–346.