

## **Supplemental Tables**

**Supplemental Table S1**

**Supplemental Table S2**

**Supplemental Table S3**

**Supplemental Table S4**

**Supplemental Table S5**

**Supplemental Table S6**

**Supplemental Table S7**

**Supplemental Table S8**

**Supplemental Table S9**

**Supplemental Table S10**

**Supplemental Table S11**

**Supplemental Table S12**

**Supplemental Table S13 (Excel sheet, uploaded separately)**

**Supplemental Table S14**

**Supplemental Table S1:** Range of quantitative traits in original collection and different core collections.

Traits	Original collection (3004 acc)		CC1 (231acc)		CC2 (300 acc)		Merged CC1 & CC2 (503 acc)		CC3 (503 + 17 = 520 acc)	
	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min
Days to 80% flowering	184	50	175	50	175	52	175	50	184	50
100 GW (gm)	5	0.98	4.6	0.98	4.6	0.98	4.6	0.98	5	0.98
Days to 1 <sup>st</sup> flowering	182	45	171	46	171	45	171	45	182	45
Grain length (mm)	12.7	4.4	12.7	4.4	12.4	4.4	12.7	4.4	12.7	4.4
Grain width (mm)	4.4	1.5	4.1	1.7	4.3	2.1	4.4	1.7	4.4	1.5
Panicle length (cm)	37	13	36	13	36	13	36	13	37	13
Seed coat color	99	10	88	10	99	10	99	10	99	10
Seedling height (cm)	74	12	71	16	74	13	74	13	74	12
Culm length (cm)	204	27	204	35	181	27	204	27	204	27
Culm number	40	5	40	6	33	5	40	5	40	5
Culm diameter (mm)	9.1	2	9.1	2	7.7	3	9.1	2	9.1	2

Highlighted traits were not picked up for their complete range in any mini-core collection (CC1, CC2 and CC3). Seventeen accessions were included in CC3 (503+17 = 520 accessions) to cover the entire range for all the traits with respect to original collection.

**Supplemental Table S2:** Assessment of mini-core collections for various evaluation indices using phenotypic data.

Core collection	MD%	VD%	VR%	CR%	H	I
CC1 (231 acc)	4.08	39.77	86	92	2.25	0.79
CC2 (300 acc)	2.8	19.78	107.68	91.1	1.98	0.77
CC3 (520 acc)	2.9	18.9	109.3	96.2	2.17	0.79

MD% (Mean difference percentage), VD% variance difference percentage, VR % (Variable rate of coefficient of variance), CR% (coincidence rate of range), H (Shannon diversity index), I (Nei's diversity index)

**Supplemental Table S3:** Distribution of accessions from different varietal groups in different mini-core collections to check proportionate representation from original collection.

<b>Core/varietal group</b>	<i>indica</i>	<b>Tropical japonica</b>	<b>Temperate japonica</b>	<i>japonica</i>	<i>aus/boro</i>	<b>Intermediate</b>	<b>Aromatic (Basmati)</b>
Original collection (3004 acc)	<b>1743</b>	<b>388</b>	<b>320</b>	<b>132</b>	<b>215</b>	<b>135</b>	<b>71</b>
CC1 (231 acc)	<b>129</b> (7.4%)	<b>15</b> (3.8%)	<b>38</b> (11.9%)	<b>14</b> (10.6%)	<b>11</b> (5.1%)	<b>19</b> (14%)	<b>5</b> (7%)
CC2 (300 acc)	<b>171</b> (9.8%)	<b>10</b> (2.5%)	<b>8</b> (2.5%)	<b>12</b> (9.1%)	<b>42</b> (19.5%)	<b>45</b> (33.3%)	<b>12</b> (16.9%)
CC3 (520 acc)	<b>295</b> (16.9%)	<b>27</b> (6.9%)	<b>44</b> (13.4%)	<b>23</b> (17.4%)	<b>55</b> (25.6%)	<b>61</b> (44.4%)	<b>15</b> (21%)

**Supplemental Table S4:** Distribution of accessions from different regional gene pools in different mini-core collections to check proportionate representation from original collection.

<b>Core/Regions</b>	<b>South-Asia</b>	<b>South East-Asia</b>	<b>China</b>	<b>Africa</b>	<b>America</b>	<b>Europe</b>	<b>East Asia</b>	<b>Oceania</b>	<b>Unknown</b>
Original collection (3004 acc)	<b>787</b>	<b>1016</b>	<b>482</b>	<b>252</b>	<b>166</b>	<b>118</b>	<b>132</b>	<b>17</b>	<b>34</b>
CC1 (231 acc)	<b>55</b> (6.9%)	<b>52</b> (5.1%)	<b>52</b> (10.8%)	<b>15</b> (5.9%)	<b>17</b> (10.2%)	<b>18</b> (15.2%)	<b>15</b> (11.4%)	<b>4</b> (23.5%)	<b>3</b> (8.8%)
CC2 (300 acc)	<b>122</b> (15.5%)	<b>70</b> (6.9%)	<b>55</b> (11.4%)	<b>23</b> (9.1%)	<b>13</b> (7.8%)	<b>2</b> (1%)	<b>8</b> (6%)	<b>1</b> (5.9%)	<b>6</b> (17.4%)
CC3 (520 acc)	<b>176</b> (22.4%)	<b>123</b> (12%)	<b>101</b> (20.95%)	<b>38</b> (15%)	<b>28</b> (16.9%)	<b>19</b> (16.1%)	<b>21</b> (15.9%)	<b>5</b> (29.4%)	<b>9</b> (25.5%)

**Supplemental Table S5:** Correlation coefficient analysis for quantitative traits in CC3 accessions.

Traits/traits	A	B	C	D	E	F	G	H	I	J	K
<b>A</b>	--										
<b>B</b>	-0.07	--									
<b>C</b>	<b>0.998</b>	-0.2	--								
<b>D</b>	0.07	0.44	-0.04	--							
<b>E</b>	-0.227	0.542	-0.27	0.021	--						
<b>F</b>	0.370	0.086	0.489	0.244	-0.025	--					
<b>G</b>	-0.095	0.051	-0.008	-0.02	0.084	-0.023	--				
<b>H</b>	0.139	0.037	0.156	0.153	-0.021	0.08	0.158	--			
<b>I</b>	0.555	0.065	0.68	0.165	0.005	0.557	0.069	0.292	--		
<b>J</b>	-0.033	-0.16	-0.27	-0.09	-0.208	-0.144	0.029	-0.13	-0.112	--	
<b>K</b>	0.426	0.097	0.614	0.112	-0.03	0.341	-0.03	0.096	0.391	-0.04	--

\*Values > 0.7 were significant, \*\*values between 0.3-0.7 were weakly significant, P < 0.001

A - days to 80% heading (DEH), B - 100 grain weight (HGW), C - days to first flower (DFF), D - grain length (GL), E - grain width (GW), F - panicle length (PL), G - secondary branches (SB), H- seedling height (SH), I - culm length (CL), J - culm number (CN) and K - culm diameter (CD)

**Supplemental Table S6:** Distribution of 3004 accessions of original rice collection in different clusters of maximum likelihood dendrogram (based on different varietal group).

Cluster/varietal group	<i>indica</i>	<i>japonica</i>	Tropical <i>japonica</i>	Temperate <i>japonica</i>	<i>aus/boro</i>	Intermediate	Aromatic (Basmati)
<b>Cluster Ia (1771 acc)</b>	1641 (92.6%)	7	27	30	30	26	10
<b>Cluster Ib (216 acc)</b>	25	1	2	7	172 (72.6%)	5	4
<b>Cluster IIa (519 acc)</b>	35	80 (15.4%)	329 (63.3%)	31	4	35	5
<b>Cluster IIb (358 acc)</b>	18	36	22	250 (69.8%)	6	25	1
<b>Cluster IIc (97 acc)</b>	9	5	6	1	3	23 (23.7%)	50 (51.5%)
<b>Un-clustered accessions (43 acc)</b>	15 (34.8%)	3	2	1	0	21 (48.8%)	1

**Supplemental Table S7:** Distribution of CC3 accessions (520) in different clusters of maximum-likelihood dendrogram of original collection of rice (3004 accessions).

Accession distribution in cluster of maximum-likelihood dendrogram of original collection (3004 accessions)	Distribution of accession from different clusters of maximum-likelihood dendrogram captured in CC3 (520 accessions)
Cluster Ia - 1771 accessions	Cluster Ia - 322 accessions (18.1%)
Cluster Ib - 216 accessions	Cluster Ib - 42 accessions (19.4%)
Cluster IIa - 519 accessions	Cluster IIa - 55 accessions (10.5%)
Cluster IIb - 358 accessions	Cluster IIb - 65 accessions (18%)
Cluster IIc - 97 accessions	Cluster IIc - 25 accessions (25.7%)
Un-clustered group - 43 accessions	Un-clustered group - 11 accessions (25%)

**Supplemental Table S8:** Distribution of different varietal group of original collection (3004 acc) in different clusters of FastSTRUCTURE analysis.

Varietal group/Cluster	FSTR CL1 (219 acc)	FSTR CL2 (522 acc)	FSTR CL3 (90 acc)	FSTR CL4 (973 acc)	FSTR CL5 (372 acc)	FSTR CL6 (323 acc)	FSTR CL7 (505 acc)
<i>indica</i> (1743 acc)	28	47	9	<b>885 (91%)</b>	26	<b>297 (92%)</b>	<b>451 (89.3%)</b>
<i>japonica</i> (132 acc)	1	<b>94 (18%)</b>	4	5	25	0	3
Temperate <i>japonica</i> (320 acc)	4	35	1	17	<b>248 (66.6%)</b>	6	9
Tropical <i>japonica</i> (388 acc)	1	<b>310 (59.3%)</b>	2	26	35	5	9
<i>aus/boro</i> (215 acc)	<b>179 (81.7%)</b>	3	5	14	5	4	5
Intermediate (135 acc)	3	<b>27</b>	<b>19 (21.1%)</b>	<b>21</b>	<b>29</b>	9	<b>27</b>
Aromatic (Basmati) (71 acc)	3	6	<b>50 (55.5%)</b>	5	4	2	1

**Supplemental Table S9:** Analysis of original collection (3004 accessions) for admixtures through population structure using FastSTRUCTURE.

<b>Pure accessions (1762 acc)</b>		<b>Admixtures (1242 acc)</b>	
<b>Structure analysis (K=7)</b>		<b>Structure analysis (K=7)</b>	
FSTR CL 1	189	FSTR CL 1	30
FSTR CL 2	330	FSTR CL 2	192
FSTR CL 3	64	FSTR CL 3	26
FSTR CL 4	591	FSTR CL 4	382
FSTR CL 5	256	FSTR CL 5	116
FSTR CL 6	148	FSTR CL 6	145
FSTR CL 7	154	FSTR CL 7	351
<b>Varietal group (K=7)</b>		<b>Varietal group (K=7)</b>	
<i>indica</i>	922	<i>indica</i>	821
<i>japonica</i>	90	<i>japonica</i>	42
Temperate <i>japonica</i>	234	Temperate <i>japonica</i>	86
Tropical <i>japonica</i>	236	Tropical <i>japonica</i>	152
<i>aus/ boro</i>	186	<i>aus/ boro</i>	29
Aromatic (Basmati)	53	Aromatic (Basmati)	18
Intermediate	41	Intermediate	94
<b>Region wise (K=7)</b>		<b>Region wise (K=7)</b>	
South Asia	497	South Asia	290
South East Asia	567	South East Asia	449
China	275	China	207
Africa	166	Africa	86
America	87	America	79
Europe	53	Europe	65
East Asia	89	East Asia	43
Oceania	9	Oceania	8
Unknown	19	Unknown	15

Accessions with  $\geq 80\%$  genome similarity were considered as pure while accessions with  $< 80\%$  shared genome were termed as admixtures. Accessions highlighted in red have around equal or more number of admixtures than pure accessions.

**Supplemental Table S10:** Distribution of CC3 accessions (520 accessions) in FastSTRUCTURE derived clusters of original collection of rice (3004 accessions).

FastStructure Clusters	Accessions from original collection	Accessions of original collection with Q value > 80% (Pure)	Accessions of original collection with Q value < 80% (Admixtures)	Accessions picked from original collection in CC3 mini-core	Accessions of CC3 with Q value > 80% (Pure)	Accessions of CC3 with Q value < 80% (Admixture)
FSTR CL 1	219	189	30	50	40	10
FSTR CL 2	522	330	192	42	23	19
FSTR CL 3	90	64	26	24	13	11
FSTR CL 4	973	591	382	185	109	76
FSTR CL 5	372	256	116	74	37	37
FSTR CL 6	323	148	145	61	28	33
FSTR CL 7	505	154	351	84	25	59

Accessions with  $\geq 80\%$  genome similarity were considered as pure while accessions with  $< 80\%$  shared genome were termed as admixtures. Accessions highlighted in red have around equal or more number of admixtures than pure accessions.

**Supplemental Table S11:** Distribution of CC3 accessions (520 acc) based on varietal groups in different clusters of FastSTRUCTURE analysis (K=7). Numbers in parentheses represents accessions.

Cluster/ Varietal group	<i>indica</i> (295)	Tropical <i>japonica</i> (27)	Temperate <i>japonica</i> (44)	<i>Japonica</i> (23)	<i>aus/boro</i> (55)	Intermediate (61)	Aromatic (Basmati) (15)
CC CL1 (73 acc)	57	0	1	1	1	12	1
CC CL2 (23 acc)	1	0	0	0	5	8	9
CC CL3 (49 acc)	6	0	1	1	41	0	0
CC CL4 (70 acc)	58	1	0	0	3	8	0
CC CL5 (43 acc)	3	17	4	10	0	8	1
CC CL6 (191 acc)	165	5	2	2	4	11	2
CC CL7 (71 acc)	5	4	36	9	1	14	2

**Supplemental Table S12:** Analysis of CC3 (520 accessions) for admixtures through population structure using FastSTRUCTURE.

<b>Pure accessions (275 accessions)</b>		<b>Admixtures (245 accessions)</b>	
<b>Structure analysis (K=7)</b>		<b>Structure analysis (K=7)</b>	
CC CL1	25	CC CL1	48
CC CL2	13	CC CL2	10
CC CL3	40	CC CL3	9
CC CL4	28	CC CL4	42
CC CL5	23	CC CL5	20
CC CL6	109	CC CL6	82
CC CL7	37	CC CL7	34
<b>Varietal group (K=7)</b>		<b>Varietal group (K=7)</b>	
<i>indica</i>	151	<i>indica</i>	144
<i>japonica</i>	11	<i>japonica</i>	12
Temperate <i>japonica</i>	29	Temperate <i>japonica</i>	15
Tropical <i>japonica</i>	14	Tropical <i>japonica</i>	13
<i>aus/ boro</i>	45	<i>aus/ boro</i>	10
Aromatic (Basmati)	11	Aromatic (Basmati)	4
Intermediate	14	Intermediate	47
<b>Region wise (K=7)</b>		<b>Region wise (K=7)</b>	
South Asia	106	South Asia	70
South East Asia	60	South East Asia	63
China	47	China	54
Africa	23	Africa	15
America	15	America	23
Europe	6	Europe	13
East Asia	11	East Asia	10
Oceania	2	Oceania	3
Unknown	5	Unknown	4

Accessions with  $\geq 80\%$  genome similarity were considered as pure while accessions with  $< 80\%$  shared genome were termed as admixtures. Accessions highlighted in red have equal or more number of admixtures than pure accessions



**Supplemental Table S14:** Association analysis of 3004 accessions of original collection.

Trait	Chr	Position	Major allele	Minor allele	Minor allele frequency	Nipp. allele	CC3 allele	P-value FDR adjusted	R <sup>2</sup> value (%)	Known loci
Grain length	3	16733441	G	T	0.36	G	T	3.4 X 10 <sup>-43</sup>	43.5	<i>GS3</i>
Grain length	5	5361894	G	A	0.36	G	A	3.4 X 10 <sup>-9</sup>	38.8	<i>qSW5</i>
Grain width	5	5371686	C	T	0.49	C	T	9.3 X 10 <sup>-34</sup>	51.4	<i>qSW5</i>
Grain width	5	28019687	T	C	0.10	T	C	8.4 X 10 <sup>-6</sup>	48	
Hundred Grain weight	3	16733441	G	T	0.36	G	T	7.9 X 10 <sup>-5</sup>	35.2	<i>GS3</i>
Hundred Grain weight	5	5375201	T	C	0.48	T	C	7.9 X 10 <sup>-5</sup>	35.2	<i>qSW5</i>
Endosperm type	6	1731808	G	C	0.20	G	C	1.03 X 10 <sup>-29</sup>	20.2	<i>waxy</i>
Endosperm type	6	6830286	G	A	0.21	G	A	3.4 X 10 <sup>-8</sup>	15.6	
Seed coat color	7	6133394	G	A	0.26	G	A	6.6 X 10 <sup>-11</sup>	7.2	<i>Rc</i>
Seed coat color	7	6417000	G	T	0.32	G	T	1.7 X 10 <sup>-10</sup>	7.1	
Seed coat color	7	6656052	T	C	0.43	T	C	1.8 X 10 <sup>-8</sup>	6.8	
Seed coat color	2	32431463	A	G	0.27	A	G	3.7 X 10 <sup>-5</sup>	5.6	
Panicle threshability	2	21739453	C	T	0.23	C	T	6.8 X 10 <sup>-3</sup>	16.4	

Nipp; Nipponbare, Chr; Chromosome