Comparative analysis of *Miscanthus* and *Saccharum* reveals a shared whole-genome duplication but different evolutionary fates

Changsoo Kim

**Recent publication in The Plant Cell**

Comprehensive Analysis of Miscanthus and Saccharum Reveals a Shared Whole-Genome Duplication but Different Evolutionary Fates


**Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia 30602**

† Genomics and Computational Biology, College of Life Sciences, Seoul National University, Seoul, South Korea

S. Korea

Recent publication in *The Plant Cell*

**Grass Phylogeny** by Piperno and Sues, 2005

Subfamilies

- **BEP clade**
  - PACCAD clade

- **BEP clade**
  - PACCAD clade

Tribe Andropogonae

**Saccharum** spp.

- a source of sugar for hundreds of years
  (being cultivated more than 20M ha in tropical- and subtropical regions)

- of increasing importance as a source of bioethanol
  (1st generation bioethanol)

**Miscanthus** spp.

- Temperate grass with a large amount of biomass
- of increasing importance as a source of lignocellulosic bioethanol
  (2nd generation bioethanol)
Basic Hypothesis

Emergence of Sorghum and Saccharum ca. 7.9 MYA (Kim et al. 2010)

Emergence of Miscanthus and Saccharum ca. 6.5 MYA (Yue et al. 2012)

Saccharum

Miscanthus

Sorghum

Two additional genome duplications after the pan-cereal duplication (Kim et al. 2008)

Sorghum-specific duplication

One additional duplication after the pan-cereal duplication (Kim et al. 2012; Yue et al. 2012)

Miscanthus-specific duplication

Evolutionary analysis

Materials and Methods

• Two NGS platforms
  – 454 GX Flex Titanium
  – Illumina HiSeq 2000

• 3 accessions of Miscanthus sinensis were used
  – Two accessions for 454 (SNP discovery also performed)
  – One accession (collected in S. Korea) for Illumina

Sequencing results from two platforms

<table>
<thead>
<tr>
<th></th>
<th>HiSeq 2000</th>
<th>454 GX Flex Titanium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read number</td>
<td>5.053 million</td>
<td>6.7 million</td>
</tr>
<tr>
<td>Read length</td>
<td>~337 Gb</td>
<td>~1.4 Gb</td>
</tr>
<tr>
<td>Contig number(^1)</td>
<td>~18.6 million</td>
<td>121,220</td>
</tr>
<tr>
<td>Contig length</td>
<td>~2.2 Gb</td>
<td>~85 Mb</td>
</tr>
<tr>
<td>N50</td>
<td>1,119</td>
<td>486</td>
</tr>
<tr>
<td>Coverage(^2)</td>
<td>~0.05 x</td>
<td>~0.016 x</td>
</tr>
</tbody>
</table>

\(^1\) Contigs > 200 bp are only considered

\(^2\) Using contig lengths based on the genome size of M. sinensis (2.6 Gb) estimated by flow cytometry

Evolutionary analysis - Illumina

Orthologs b/w Sb-M

Orthologs & paralog signals

Evolutionary analysis

Frequencies of internal trees when using sorghum or Saccharum

<table>
<thead>
<tr>
<th></th>
<th>Sb</th>
<th>So</th>
<th>Sb &lt; So(^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of trees tested</td>
<td>29,326</td>
<td>40,009</td>
<td></td>
</tr>
<tr>
<td>Number of internal trees</td>
<td>13,389</td>
<td>20,982</td>
<td></td>
</tr>
<tr>
<td>Internal tree frequency</td>
<td>45.7%</td>
<td>52.4%</td>
<td>***</td>
</tr>
</tbody>
</table>
Combined evolutionary structure of Saccharinae subtribe

Origin of the shared duplication event

- The shared genome duplication might trigger the divergence of genera *Miscanthus* and *Saccharum*.
- *Miscanthus* has disomic inheritance whereas *Saccharum* has polysomic inheritance.
- The question becomes how major genomic events have caused different reproductive patterns in those two close genera.

Origin of the shared duplication event

**Case I. Autotetraploidization**

In this case, paralogs are not diverging or slowly diverging due to extensive genetic material exchanges among them.

Origin of the shared duplication event

**Case II. Allotetraploidization**

In this case, paralogs are diverging between subgenomes because no genetic material exchange occurs.

Origin of the shared duplication event

Our Ks profile is quite consistent with allotetraploidization.

Origin of the shared duplication event

The shared allotetraploidization might cause disomic inheritance patterns but an additional autopolyploidization in *Saccharum* might switch its disomic pattern to polysomic pattern.
Additional suggestion in the variation of basal chromosome numbers in *Saccharum* spp.

- Universal tendency after the whole genome doubling
  - Chromosome numbers return to a relatively narrow range, as exemplified in *Arabidopsis* (only 5 gametic chromosomes)

- The basal chromosome numbers of *Saccharinae* and *Sorghinae* subtribes are thought to have \( x = 10 \).

- Three representative species in *Saccharum* spp.
  - *S. officinarum* (\( x = 10 \) mostly)
  - *S. robustum* (\( x = 10 \) mostly)
  - *S. spontaneum* (\( x = 8 \) but some variations in the basal chromosome numbers)

Additional suggestion in the variation of basal chromosome numbers in *Saccharum* spp. (Data adapted from Irvine 1999)

Take-home messages

- *Miscanthus* and *Saccharum* share the whole genome allotetraploidization after their divergence from the sorghum lineage.

- *Miscanthus* has maintained disomic inheritance due to the allotetraploidization whereas an additional autotetraploidization specific to *Saccharum* changed its disomic inheritance to polysomic inheritance but two different polyploidizations caused two pseudo-paralogous groups of chromosomes.

Acknowledgement

- INTERNATIONAL CONSORTIUM FOR SUGARCANE BIOTECHNOLOGY
  - James E. Irvine Memorial Travel Award

- PGML
  - Dr. Andrew Paterson (Correspondence)
  - Dr. Xiyin Wang (Co-first)
  - Dr. Tae Ho Lee

- Chungnam National University, South Korea
  - Dr. Geung-Joo Lee

- Mendel biotechnology
  - Dr. Katrin Jakob