Developing a Genomic Toolbox for the Improvement of Intermediate Wheatgrass as a Perennial Grain Crop

Kevin Dorn
Postdoctoral Fellow
Kansas State University
Poland Lab

This talk is Twitter friendly!
@Dornomics

Intermediate Wheatgrass (Thinopyrum intermedium)

Perennial grass species targeted for direct domestication into grain/biomass crop

> 1,000 kg/ha grain yield & biomass yields comparable to switchgrass

deep root system = ecosystem services

Domesticating Intermediate Wheatgrass

- Significant progress has been made breeding for key traits
- Doubled grain yield & seed size in last decade
- Challenges still exist - consistent yields - consistent maturity - free threshing

Intermediate Wheatgrass Products = Kernza™

The Land Institute trademarked the processed grain from IWG as ‘Kernza’

Significant commercial interest for perennial grain crops

The challenge with intermediate wheatgrass...

- Segmental Allohexaploid (unclear progenitors)
- 2n=6x=42
- 1C=12.6 Gb
- Obligate outcrosser
Developing a Genomic Toolbox for IWG

1.) Sequencing, assembly, validation, anchoring/ordering
2.) First applications for the IWG genome

Genome sequencing

Haploid plant derived from twin seedling by The Land Institute

Genome assembly

<table>
<thead>
<tr>
<th>Number</th>
<th>Scaffold</th>
<th>Contig</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>237,622</td>
<td>1,059,461</td>
</tr>
<tr>
<td>Total Assembly Length (bp)</td>
<td>11,605,214,228 (92%)</td>
<td>11,193,042,853</td>
</tr>
</tbody>
</table>

Genome assembly – BUSCO analyses

| Number of scaffolds | 237,622 |
| Total assembly length (bp) | 11,605,214,228 (92%) |

Complete BUSCOs | 917 (95.92%)
Copy Number | # BUSCOs
1 | 37
2 | 71
3 | 654
4 | 155

BUSCO = universally conserved single-copy gene
Population Sequencing (POPSEQ)

Two F1 populations
- C4-5353 x C4-8134 (92 F1 individuals)
- C4-5353 x C4-2856 (92 F1 individuals)

POPSEQ Marker Generation

- Parents In Cross
  - C4-2856
  - C4-5353
- Align to C4-5353 Assembly
- Call SNPs
- C4-5353 unique 75mer Markers

Roadmap to V1.0 genome

- HIC Scaffolding
- Scaffolded NRGene Assembly
- Identify Mismatches
- Integrated Assembly
- Polished Assembly
- Illumina Fragment Data (eliminate homozygous variants)

A total of 2.3M markers have been generated.

- ~700 breaks identified, or a break every 16 Mb
- ~10Gb have at least 25 markers for use in ordering
Developing a Genomic Toolbox for IWG

1.) Sequencing, assembly, validation, anchoring/ordering
2.) First applications for the IWG genome

2.) Identify progenitor species of IWG
- Many conflicting studies on evolution of IWG genome and progenitors
  Most based on minimal markers / GISH studies
  Best current thought: Thinopyrum elongatum (J) - Pseudoroegneria spp. (S)
  Thinopyrum bessarabicum (J')
  Other diploid species suggested as progenitors:
  Dasypyrum villosum (V)
  Taeniatherum spp. (Ta)
  Aegilops tauschii (D)

Genotype-By-Sequencing of ~60 potential progenitor species
Dataset courtesy of Xiaofei Zhang

Extract 64-mers shared within each potential progenitor species
Align to IWG pseudochromosomes
Utilizing the assembly: Tools for the breeding program

Major domestication traits: shattering & free threshing seed

Wide variability of these traits in breeding populations

Major QTLs for these traits in intermediate wheatgrass

Goal: direct marker for causal variant(s)

Q — cloned in bread wheat
Tenacious Glume — not cloned
Brittle rachis — cloned in barley

Summary

1.) Towards chromosome scale assembly of the IWG genome
- NBGene genome assembly
- 10X Linked-Reads scaffolding
- Initial anchoring to GBS consensus map
- POPSEQ anchoring ongoing
- HIC scaffolding
- Reference transcriptome for annotation

2.) Utilizing the assembly
- First genome-wide sequence based analysis of progenitors
- New tools for breeding programs

Acknowledgements

Kansas State University
Jesse Poland
Traci Kantariski
Jared Crain
Mark Lucas
The Land Institute
Lee Dehnan
Shuwen Wang
Kathryn Turner
HudsonAlpha
Jeremy Schmutz
Jane Glenwood
Chris Pruitt
Jerry Jenkins
USDA – Logan, Utah
Steve Lander
University of Minnesota
Jim Anderson
Xiaoxue Zhang
Kayla Allendorf
Don Wyse
David Marks

Funding Sources

Department of Energy – Joint Genomes Institute Community Science Program

The Land Institute
Perennial Agriculture Project & Malone Foundation

Minnesota Dept. of Agriculture

UMN Forever Green Initiative

Kevin Dorn  @Dornomics  www.kdorn.com