Molecular ideotype breeding and rice quality

Dan Waters
Research in rice grain eating quality

Genotype x Environment = Quality

Genotyping

Phenotyping

Genotype x Environment = Human Perception = Quality

International Temperate Rice Conference
March 6-9, 2017
Grain quality ideotype

- Both yield and grain quality controlled by an interacting (epistatic) gene network

- Rebuilding both yield and grain quality gene networks is challenging

- The relative importance of each component of the grain quality network is (genetic) background (and environment) dependent -> the molecular composition of high quality cultivars provide the template or ideotype
The rice grain

Oryza sativa cv Nipponbare

~ 90 - 95% Starch
~ 4 - 10% Protein
~ 1% Lipid
<1% Secondary metabolites

International Temperate Rice Conference
March 6-9, 2017
Grain starch synthesis genes are known

AGPase

Granule bound starch synthase (GBSS)
GBSSI

Starch synthase (SS) family
SSI, SSIIa, SSIIIa, SSIIIb, SSIVa, SSIVb

Starch Branching enzymes (SBES)
BEI and BEIIb

Debranching Enzymes (DBEs)
ISA1, ISA2, Pullulanase
Association of starch genes and traits I

- 33 indica and 37 japonica cultivars
- Majority of starch genes influence GT, AC and GC
- This network of eating quality control dominated by GBSSI and SSIII-3 (SSIIa)

Tian et al PNAS 106:21760-21765
Association of starch genes and traits II

• Association of starch biosynthesis genes in **233 Australian** rice breeding lines with RVA, apparent amylose content (AAC), gelatinisation temperature (GT) and retrogradation

  • GBSSI and BEI; AAC, retrogradation and RVA parameters

  • SSIIa, SSIIIa, SSIIIb and SSIVa; GT and RVA parameters

  • SSI, ISA2 and Pullulanase had small but measurable impact on starch traits

  • SSIVb and ISA1, no sequence variation and therefore no impact on starch traits
  
  -> may vary and play a role in other germplasm

• Only BEIIb displayed sequence variation and no association with starch traits

Kharabian-Masouleh
Scientific Reports 2: 557
Association of starch genes and traits III

Mo et al.,
Breeding Sci.
64: 164–175

- **187 (170 + 17)**
  - Korean breeding lines
- **GBSSI fixed in both populations**
- Many gene-trait associations disappear in the larger, but more narrow, population

### Table 5. Summary of association analyses ($R^2$) between SSRG marker alleles and grain quality traits among 187 non-glutinous Korean rice varieties under different temperature conditions

<table>
<thead>
<tr>
<th>Gene (Marker)</th>
<th>Trait</th>
<th>All ($n = 187$)</th>
<th>Japonica ($n = 170$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>G1</td>
<td>GII</td>
</tr>
<tr>
<td><strong>AGPase-L (InDel)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CGP</td>
<td>0.021*</td>
<td>0.078***</td>
<td>0.136***</td>
</tr>
<tr>
<td>AC</td>
<td>0.023*</td>
<td>0.057***</td>
<td>0.024*</td>
</tr>
<tr>
<td>ADV</td>
<td>0.045**</td>
<td>0.038**</td>
<td>0.069***</td>
</tr>
<tr>
<td>PKV</td>
<td>0.106***</td>
<td>0.291***</td>
<td>0.131***</td>
</tr>
<tr>
<td>SBV</td>
<td>0.315***</td>
<td>0.313***</td>
<td>0.297***</td>
</tr>
<tr>
<td>BDV</td>
<td>0.432***</td>
<td>0.418***</td>
<td>0.344***</td>
</tr>
<tr>
<td><strong>AGPase-S (C/T)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBV</td>
<td>0.074***</td>
<td>0.035*</td>
<td>0.034*</td>
</tr>
<tr>
<td>BDV</td>
<td>0.081***</td>
<td>0.104***</td>
<td>0.074***</td>
</tr>
<tr>
<td><strong>AGPase-S (A/T)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBV</td>
<td>0.152***</td>
<td>0.123***</td>
<td>0.119***</td>
</tr>
<tr>
<td>BDV</td>
<td>0.178***</td>
<td>0.144***</td>
<td>0.134***</td>
</tr>
<tr>
<td><strong>SSIIa (TT/GC)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADV</td>
<td>0.297***</td>
<td>0.204***</td>
<td>0.141***</td>
</tr>
<tr>
<td>PT</td>
<td>0.110***</td>
<td>0.043***</td>
<td>0.050***</td>
</tr>
<tr>
<td><strong>SSB1 (InDel)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADV</td>
<td>0.078***</td>
<td>0.049***</td>
<td>0.091***</td>
</tr>
<tr>
<td>PKV</td>
<td>0.135***</td>
<td>0.188***</td>
<td>0.070***</td>
</tr>
<tr>
<td>SBV</td>
<td>0.215***</td>
<td>0.205***</td>
<td>0.196***</td>
</tr>
<tr>
<td>BDV</td>
<td>0.312***</td>
<td>0.286***</td>
<td>0.236***</td>
</tr>
<tr>
<td><strong>SSB1II (C/G)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKV</td>
<td>0.156***</td>
<td>0.224***</td>
<td>0.100***</td>
</tr>
<tr>
<td>SBV</td>
<td>0.284***</td>
<td>0.275***</td>
<td>0.261***</td>
</tr>
<tr>
<td>BDV</td>
<td>0.377***</td>
<td>0.332***</td>
<td>0.392***</td>
</tr>
<tr>
<td><strong>SDB1 (InDel)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADV</td>
<td>0.056*</td>
<td>0.063***</td>
<td>0.033*</td>
</tr>
<tr>
<td>PKV</td>
<td>0.076***</td>
<td>0.110***</td>
<td>0.110***</td>
</tr>
<tr>
<td>SBV</td>
<td>0.157***</td>
<td>0.238***</td>
<td>0.127***</td>
</tr>
<tr>
<td>BDV</td>
<td>0.225***</td>
<td>0.214***</td>
<td>0.212***</td>
</tr>
<tr>
<td><strong>SDB1 (InDel)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

International Temperate Rice Conference
March 6-9, 2017
Starch quality ideotype

• The relative importance of each starch synthesis gene/allele is genetic background/breeding program dependent

• Current DNA sequencing technology allows individual breeding programs to track all starch synthesis genes and alleles within their germplasm to re-build starch synthesis gene networks
And rice grain proteins?

 Defined by solubility differences

• PB-II
  → globulins, salt soluble
  ~5-10% of rice grain, ~5% of wheat grain

  → glutelins, dilute acid/alkali soluble
  ~65-80% of rice grain, ~40% of wheat grain

• PB-I
  → prolamins, alcohol soluble, poorly digested
  ~10-20% of rice grain, ~55% of wheat grain

Ashida et al.
Breeding Science
61: 201-207

International Temperate Rice Conference
March 6-9, 2017
Rice grain proteins

Protein bodies

Starch granules

Kasem et al.
Rice 4: 12-20

International Temperate Rice Conference
March 6-9, 2017
Rice grain HPLC prolamin profiles

- Qualitative and quantitative protein composition differences
Long and medium grain HPLC profiles and texture

- HPLC protein profiles of **160 advanced breeding lines** (80 long grain and 80 medium grain) 17% -20% amylose correlated (≤-0.5 or ≥0.5) with texture parameters

- Albumin and globulin display very low levels of variation

<table>
<thead>
<tr>
<th></th>
<th>Long grain Profile 1</th>
<th>Long grain Profile 2</th>
<th>Medium grain Profile 1</th>
<th>Medium grain Profile 2</th>
</tr>
</thead>
<tbody>
<tr>
<td># Samples</td>
<td>53</td>
<td>27</td>
<td>67</td>
<td>13</td>
</tr>
<tr>
<td>% Protein</td>
<td>6.3 - 10.3</td>
<td>6.4 - 8.5</td>
<td>5.6 - 7.8</td>
<td>5.5 - 9.8</td>
</tr>
<tr>
<td>Pasting Temp (°C)</td>
<td>68 – 77</td>
<td>67–78</td>
<td>66 – 69</td>
<td>66 –76</td>
</tr>
<tr>
<td>RVA correlations</td>
<td>Prolamin</td>
<td>NS</td>
<td>Prolamin</td>
<td>Protein (prolamin)</td>
</tr>
<tr>
<td>TA correlations</td>
<td>Glutelin</td>
<td>NS</td>
<td>Prolamin</td>
<td>Glutelin</td>
</tr>
</tbody>
</table>

International Temperate Rice Conference
March 6-9, 2017
Medium grain samples

- 80 samples of amylose content 17 - 20%

• Profile 1; **67 samples**
  - “Typical” medium grain profile
  - 5.6% - 7.8% protein,
  - Pasting Temp 66 – 69°C

• Profile 2; **13 samples**
  - “New” profile
  - 5.5% - 9.8% protein
  - Pasting temp 66 – 76°C
Medium grain Profile 1

• 67 samples; prolamsins most highly correlated (≤-0.5 or ≥0.5) with texture parameters

<table>
<thead>
<tr>
<th>Texture parameter</th>
<th>Correlation Distribution</th>
<th>Total Prolamin Correlation</th>
<th>Total Protein Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA Height</td>
<td>18/23 of peaks 0.57 to 0.85</td>
<td>0.83</td>
<td>0.57</td>
</tr>
<tr>
<td>RVA Setback</td>
<td>15/23 of peaks 0.50 to 0.74</td>
<td>0.66</td>
<td>0.42</td>
</tr>
<tr>
<td>TA Custom Force Distance</td>
<td>15/23 of peaks 0.5 to 0.64</td>
<td>0.63</td>
<td>0.45</td>
</tr>
<tr>
<td>RVA Trough 1</td>
<td>14/23 of peaks -0.51 to -0.81</td>
<td>-0.67</td>
<td>-0.31</td>
</tr>
<tr>
<td>RVA Breakdown 1</td>
<td>17/23 of peaks -0.51 to -0.73</td>
<td>-0.72</td>
<td>-0.54</td>
</tr>
<tr>
<td>RVA Peak 1</td>
<td>20/23 of peaks -0.5 to -0.79</td>
<td>-0.81</td>
<td>-0.51</td>
</tr>
<tr>
<td>Millout</td>
<td>9/23 of peaks 0.5 to 0.63</td>
<td>0.56</td>
<td>0.39</td>
</tr>
</tbody>
</table>

• “Typical” profile; pattern of correlations very similar to “Sushi” samples

• 5.6% - 7.8% protein, Pasting Temp 66 – 69°C

• Amylose and globulins weak correlations

• Glutelin correlations of opposite sign to prolamin correlations
  -> Prolamin/Glutelin ratio?
Protein composition distribution

- Continuous distribution of Prolamin / Prolamin + Glutelin ratio -> many genes of small effect?
- Prolamin / Prolamin + Glutelin ratio a better predictor of Setback (cooked rice firmness) than Total Protein
Protein composition inheritance

- Protein, albumin, globulin, glutelin and prolamin display a (semi) continuous distribution around a mean -> several genes of small effect?

- Are the several genes of small effect the protein coding genes themselves?

- The gene sequence of globulins, glutelins and prolamins are known

Zhang et al.,
J. Cereal Sci.
48: 539-547
Prolamin profile inheritance

- Protein profiles may have non-continuous distribution -> few gene/s of large effect?

Huebner et al.,
Cereal Chem. 68: 583-588

**Table I**

<table>
<thead>
<tr>
<th>Variety</th>
<th>Selection</th>
<th>Pedigree</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR28</td>
<td>IR2061-214-3-8-2</td>
<td>IR833-6-1-1-1/IR1561-149-1/IR1737</td>
</tr>
<tr>
<td>IR29</td>
<td>IR2061-464-4-14-1</td>
<td>IR20*2/O. nivara/CR94-13</td>
</tr>
<tr>
<td>IR32</td>
<td>IR2070-747-6-3-2</td>
<td>IR36</td>
</tr>
<tr>
<td>IR38</td>
<td>IR2070-423-2-5-6</td>
<td></td>
</tr>
<tr>
<td>IR40</td>
<td>IR2070-414-3-9</td>
<td></td>
</tr>
<tr>
<td>IR36</td>
<td>IR2071-625-1-252</td>
<td>IR1561-228-1-2/IR1737/CR94-13</td>
</tr>
<tr>
<td>IR42</td>
<td>IR2071-586-5-6-3-4</td>
<td></td>
</tr>
</tbody>
</table>

*Varieties developed by the International Rice Research Institute.

bFemale and male parents are separated by a slash to indicate a single cross and by two slashes (///) to indicate the second cross. Asterisk (*) denotes backcrosses, and the number in the superscript indicates the dosage of the recurrent parent.
To summarise

• Population structure/selection history (and environment) influence trait components

• Complexity of grain quality network interaction is simplified by recreating molecular ideotype of high quality cultivars

• Current understanding (and technology) allows for grain quality ideotype consisting of all starch structural genes and protein composition

• Improved knowledge of protein composition inheritance may allow construction of early generation markers for protein composition
Acknowledgements

Jeanette Balindong, Ben Liu, Laura Pallas, Rachelle Ward, Margrit Martin, Ben Ovenden, Peter Snell, Robert Henry, Russell Reinke, Ardashir Kharabian-Masouleh

NSWDPI rice breeding and quality evaluation program
Method development - grain types differ by HPLC profile

- Solubility differences likely more significant than size differences -> HPLC
- HPLC quantitative -> allows statistics
Marker assisted backcrossing

'TARGET GENE' (‘FOREGROUND’) SELECTION

'TARGET GENE'

Target gene

'RECOMBINANT' SELECTION

'BACKGROUND' SELECTION

MacKil et al.,
Fifth International Rice Genetics Symposium,
Manila, 2005

International Temperate Rice Conference
March 6-9, 2017
Pushing the background forward

If backcrossing to the better quality parent, including the maximum number of grain quality genes from the better quality parent facilitates rebuilding grain quality genetic networks.
Protein influences cooking quality

Flour with (bold line) and without protein

Journal of Cereal Science
36:285-294

Southern Cross University,
July 25, 2012