Conserving and mining plant genetic resources using knowledge of the genome

Theo van Hintum & Rob van Treuren
Centre for Genetic Resources, The Netherlands (CGN)
Conserving and mining PGR

- introduction
  - PGR management
  - new approaches and tools
- molecular marker applications in PGR management
  - detailed case study (apple)
  - brief snapshots of selected CGN research
- Generation Challenge Programme
  - genomic approaches to mining PGR
- conclusions
### Conserving and mining PGR

- **plant genetic resources (PGR) conservation evolved from classical taxonomy: qualitative approach**
  - number of accessions rather than diversity captured
  - ‘having it’ more important than ‘knowing it’

- **State of the World (FAO 1996)**
  - 6.1 million accessions in 1308 collections
  - major crops
    - wheat: 774,500 accessions
    - barley: 485,000 accessions
    - rice: 408,500 accessions
    - maize: 277,000 accessions
Conserving and mining PGR

- current trend *ex situ* PGR management
  - from qualitative
    - collecting, documenting, evaluating individual accessions
  - to quantitative
    - optimising genebank management strategies by approaching the collection as a whole
    - screening large parts of the collections with genomics based techniques
Conserving and mining PGR

- new approaches
  - improving genebank efficiency and quality
    - genebank economy
    - genebank quality management
    - accession management
  - stimulating use
    - web interfaces
    - search strategies
    - identifying new alleles
Conserving and mining PGR

- **new tools**
  - molecular genetics
    - marker technology
    - allele mining techniques
  - information and communication technology
    - Internet
    - bioinformatics
Conserving and mining PGR

- **application molecular markers in PGR management**
  - acquisition
    - improving sampling strategies
    - identifying gaps and improving composition
  - maintenance
    - measuring and reducing genetic drift/shift
    - observing contamination
  - characterization
    - fingerprinting / diversity studies
  - utilization
    - creation of core collections
    - allele mining
Conserving and mining PGR

- examples use molecular markers at CGN
  - support Dutch apple NGO’s
    - detailed description
  - various other examples CGN studies using molecular markers to improve PGR management
    - lettuce
    - potato
    - white cabbage
    - barley
    - ryegrass
    - flax
    - Dutch grasslands
    - enset
Conserving and mining PGR

- molecular study of Dutch apple collections
  - total ~2000 accessions
Conserving and mining PGR

- molecular study of Dutch apple collections
  - interest CGN
      - to improve knowledge of diversity in genetic resources of apple in the Netherlands
      - to determine the degree of representation of CGNs collection
      - to improve efficiency of conservation by supporting actors
Conserving and mining PGR

- molecular study of Dutch apple collections
  - study material for microsatellite analysis
    - entire collection of CGN
    - selection from 8 other collections
      - proper documentation
      - Dutch origin
      - foreign material that has been important for cultivation in the Netherlands
    - total sample size: ~700 trees
Conserving and mining PGR

- molecular study of Dutch apple collections
  - selection of apple microsatellite markers
    - literature: 150-200 microsatellites
    - selection criteria
      - technical qualities
      - degree of variability
      - genome coverage
      - proximity to important genes
      - linkage of data to other studies
    - 16 markers selected
      - 11 of the 17 chromosomes covered
      - traits: scab resistance, canker resistance, aphid resistance, firmness, malic acid, allergy genes
Conserving and mining PGR

- molecular study of Dutch apple collections
  - identification of microsatellite variation
Conserving and mining PGR

- molecular study of Dutch apple collections
  - resolution of the microsatellite markers
    - observed number of alleles
      - 3 - 37 alleles per microsatellite
      - 296 alleles for the total set of 16 microsatellites
      - 18.5 alleles on average per microsatellite
  - probability to find identical genotypes by chance
    - depending on frequencies of the alleles in total sample
    - estimation: $6.4 \times 10^{-20} = 0.000000000000000000064$
molecular study of Dutch apple collections

- potential duplicates
  - duplication groups based on passport data
    - number of groups: 138
    - total number of trees: 355
  - verification with microsatellite data
    - duplicates confirmed: 147 (41%)
    - duplicates rejected: 202 (57%)
    - to be determined: 6
- majority of potential duplicates are different despite similarity in accession name
Conserving and mining PGR

- molecular study of Dutch apple collections
  - example of confirmed duplicates

<table>
<thead>
<tr>
<th>Variety name</th>
<th>POM-ABG</th>
<th>CGN</th>
<th>FRU/NPV</th>
<th>KRU/BAA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marker 1</td>
<td>151</td>
<td>159</td>
<td>202</td>
<td>151</td>
</tr>
<tr>
<td>Marker 2</td>
<td>156</td>
<td>166</td>
<td>161</td>
<td>159</td>
</tr>
<tr>
<td>Marker 3</td>
<td>242</td>
<td>254</td>
<td>254</td>
<td>254</td>
</tr>
<tr>
<td>Marker 4</td>
<td>147</td>
<td>176</td>
<td>176</td>
<td>176</td>
</tr>
<tr>
<td>Marker 5</td>
<td>160</td>
<td>162</td>
<td>162</td>
<td>162</td>
</tr>
<tr>
<td>Marker 6</td>
<td>155</td>
<td>161</td>
<td>161</td>
<td>161</td>
</tr>
<tr>
<td>Marker 7</td>
<td>164</td>
<td>186</td>
<td>186</td>
<td>186</td>
</tr>
<tr>
<td>Marker 8</td>
<td>95</td>
<td>103</td>
<td>103</td>
<td>103</td>
</tr>
<tr>
<td>Marker 9</td>
<td>170</td>
<td>178</td>
<td>178</td>
<td>178</td>
</tr>
<tr>
<td>Marker 10</td>
<td>123</td>
<td>137</td>
<td>137</td>
<td>137</td>
</tr>
<tr>
<td>Marker 11</td>
<td>218</td>
<td>262</td>
<td>262</td>
<td>262</td>
</tr>
<tr>
<td>Marker 12</td>
<td>178</td>
<td>182</td>
<td>182</td>
<td>182</td>
</tr>
<tr>
<td>Marker 13</td>
<td>220</td>
<td>254</td>
<td>254</td>
<td>254</td>
</tr>
<tr>
<td>Marker 14</td>
<td>207</td>
<td>209</td>
<td>209</td>
<td>209</td>
</tr>
<tr>
<td>Marker 15</td>
<td>210</td>
<td>210</td>
<td>210</td>
<td>210</td>
</tr>
<tr>
<td>Marker 16</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
</tr>
</tbody>
</table>
molecular study of Dutch apple collections
- potential duplicates based on marker data
  - total sample: 695 trees
    - duplicate groups: 119
    - number of trees involved: 339
    - redundancy: 220 (32%)
    - observed genotypes: 475 (68%)
- unexpected duplicates
  - 66% of the observed groups
- majority of identical microsatellite genotypes are not in line with the passport data
Conserving and mining PGR

- molecular study of Dutch apple collections
  - example unexpected duplication group
    - CGN Reinette Rouge Etoilée
    - ROSSEL Sterappel
    - POM-JON Paradijszoet
    - KRU/BAA Dubbele ster
    - FRU/NPV Duchesse's Favourite
Conserving and mining PGR

- molecular study of Dutch apple collections
  - potential redundancy per collection
    - potential redundancies (220)
      - within collections 73 (33%)
      - between collections 147 (67%)
    - majority of redundancies are between collections

<table>
<thead>
<tr>
<th>Collection</th>
<th>FRU/NPV</th>
<th>CGN</th>
<th>KRU/BAA</th>
<th>ROSSEL</th>
<th>POM-BEE</th>
<th>POM-JON</th>
<th>FRIJNS</th>
<th>POM-ABD</th>
<th>Total set</th>
</tr>
</thead>
<tbody>
<tr>
<td># Analyzed trees</td>
<td>241</td>
<td>175</td>
<td>80</td>
<td>78</td>
<td>60</td>
<td>30</td>
<td>16</td>
<td>15</td>
<td>695</td>
</tr>
<tr>
<td># Internal redundancies</td>
<td>36</td>
<td>19</td>
<td>3</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>220</td>
</tr>
<tr>
<td>% Internal redundancies</td>
<td>14.9</td>
<td>10.9</td>
<td>3.8</td>
<td>12.8</td>
<td>0.0</td>
<td>0.0</td>
<td>6.3</td>
<td>26.7</td>
<td>31.7</td>
</tr>
</tbody>
</table>
Conserving and mining PGR

- molecular study of Dutch apple collections
  - unique diversity within collections
    - each collection consists for about half of unique genotypes
    - majority of the alleles occurs in multiple collections

<table>
<thead>
<tr>
<th>Collection</th>
<th>FRU/NPV</th>
<th>CGN</th>
<th>KRU/BAA</th>
<th>ROssel</th>
<th>POM-BEE</th>
<th>POM-JON</th>
<th>FRIJNS</th>
<th>POM-ABD</th>
<th>Total set</th>
</tr>
</thead>
<tbody>
<tr>
<td># Analyzed trees</td>
<td>241</td>
<td>175</td>
<td>80</td>
<td>78</td>
<td>60</td>
<td>30</td>
<td>16</td>
<td>15</td>
<td>695</td>
</tr>
<tr>
<td># Unique genotypes</td>
<td>125</td>
<td>104</td>
<td>38</td>
<td>38</td>
<td>30</td>
<td>12</td>
<td>8</td>
<td>1</td>
<td>356</td>
</tr>
<tr>
<td># Observed alleles</td>
<td>210</td>
<td>237</td>
<td>203</td>
<td>175</td>
<td>149</td>
<td>134</td>
<td>122</td>
<td>97</td>
<td>296</td>
</tr>
<tr>
<td># Unique alleles</td>
<td>13</td>
<td>37</td>
<td>17</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>78</td>
</tr>
</tbody>
</table>

Centre for Genetic Resources, the Netherlands
Conserving and mining PGR

- molecular study of Dutch apple collections
  - management implications apple study
    - improvement of documentation
    - improvement of variety determination
    - identification of redundancies and removal of internal duplicates
    - identification of unique material (guidelines for safety duplication)
    - formation of a network of co-operating collections
Conserving and mining PGR

- optimization of protocol for adding modern lettuce varieties to CGN collection
  - comparison new varieties with CGN collection
  - information from breeders
  - microsatellite study
  - outcome: improved protocol to select varieties with highest added value

Conserving and mining PGR

- reduction size CGN’s potato collection
  - about 2000 accessions
    - 110 wild species
  - 314 accessions of the series Acaulia
  - AFLP study
  - taxonomic validation
  - sampling strategies
  - outcome: considerable rationalization and optimisation collection strategy

Conserving and mining PGR

- study of diversity and validation regeneration protocol in CGN cabbage collection
  - comparison drift/shift during regenerations with distances between accessions
  - sampling after ‘real’ regeneration
  - AFLP study
  - outcome: need to improve regeneration protocol or collection composition

Conserving and mining PGR

- intra-accession variation in barley
  - improvement of regeneration methods
    - avoidance of genetic drift
    - avoidance of selection
  - case study in barley
    - cultivars, landraces and wild material
  - AFLP study
  - outcome: methodological knowledge and recommendations for splitting accessions

Conserving and mining PGR

- regeneration of perennial ryegrass
  - variation in pollination
  - loss of genetic integrity
  - paternity analysis
  - microsatellite study
  - outcome: methodological knowledge and recommendations for regeneration protocol improvement

Conserving and mining PGR

- redundancy study in flax
  - CGN collection: ~1000 accessions
    - 317 accessions of breeder’s lines
    - many similar accession names
    - poorly documented
    - hardly requested for utilisation
  - AFLP study
  - outcome: recommendations for bulking accessions

Treuren et al. 2001, TAG 103: 144-152.

Centre for Genetic Resources, the Netherlands
Conserving and mining PGR

- development *in situ* conservation strategy for old Dutch grasslands
  - perennial ryegrass, white clover and Kentucky bluegrass
  - comparison of old grasslands to nature reserves and cultivars
  - AFLPs and morphological variation
  - outcome: support to policy decisions for *in situ* conservation

Conserving and mining PGR

- study clonal variation in Ethiopian enset
  - on farm conservation in Ethiopia
  - many synonyms
  - AFLP and morphological study
  - identification of duplicates
  - outcome: rationalisation conservation program

Conserving and mining PGR

- **Generation Challenge Programme**
  - an international initiative to use genomic science to mine PGR collections of major crops and use it in poor framers crops
  - initiative of the CGIAR
Conserving and mining PGR

Generation Challenge Programme

objectives

• utilize plant genetic diversity for crop improvement
  – develop low cost technologies for germplasm genotyping, allele mining and molecular breeding
  – integrate genomics with plant physiology and/or biology
• capacity building and delivery of products
  – create and enable adoption of mechanisms for products to reach end-users

Centre for Genetic Resources, the Netherlands
Conserving and mining PGR

- Generation Challenge Programme
  - consortium

Centre for Genetic Resources, the Netherlands
Conserving and mining PGR

- **Generation Challenge Programme**
  - **some facts**
    - launched in August 2003 (10 year framework two 5 year periods)
    - target areas: marginal environments in Sub-Sahara Africa, Latin America and South East Asia
    - mandate crops: all CGIAR crops
    - mandate traits: drought tolerance and associated traits
    - budget c. $16M / year of which 70% for research
    - major donors
      - European Union
      - UK Department for International Development
      - Bill and Melinda Gates Foundation
Conserving and mining PGR

Generation Challenge Programme
- research structure

- **SP1** germplasm
- **SP2** genes
- **SP3** MAS
- **SP4** data & methods
- **SP5** training / capacity

breeding programmes
improved genotypes

improved germplasm in farmers’ fields

Centre for Genetic Resources, the Netherlands
Conserving and mining PGR

- Generation Challenge Programme
  - genomics: central component
    - large scale fingerprinting of all crops
      - up to 50 microsatellites x 3000 accessions
      - development hi-throughput technologies / pipelines
    - candidate gene identification
      - forward and reverse genetics
      - QTL analysis
    - allele mining
      - EcoTILLING
      - ‘mining strategies’
    - study of orthologous candidate genes
Conserving and mining PGR

- Generation Challenge Programme
  - expected outputs relevant to PGR management
    - knowledge about the diversity in germplasm collections (and methodologies to study this)
    - methods and tools for identifying allelic diversity in germplasm collections based on DNA sequences
Conserving and mining PGR

- conclusions
  - molecular marker technology has great potential to improve PGR management
    - but little impact so far
  - applications are aimed at specific research questions, not at routine operations
  - costs of molecular analysis are still limiting factor in most cases
    - high throughput technology is required