Genes for seed quality
A Physiological Genetical Genomics approach
in tomato and Arabidopsis

Wilco Ligterink
Ronny Joosen
Seed Quality – a complex trait

- High quality seeds are undamaged seeds that have a high level of germination, which will produce uniform, vigorous seedlings without defects under various environmental conditions (Dickson 1980)
Seed Quality determined by:

**genotype**
- maternal
- paternal
- embryonic

**environment during development and maturation**
- day length
- light quality
- temperature
- soil moisture
- carbon dioxide levels
- competition
- fungi
- herbicides
- length of growing season
- mineral nutrition
- physiological age
- position on mother plant
- defoliation
- etc.
Objectives

- Molecular dissection of important processes related to seed and seedling performance
  - Development of molecular markers to aid in marker assisted breeding
  - Enable monitoring and prediction of seed quality during production and processing
  - Genetic modification to enhance seed quality
Seed quality traits are polygenic and many different processes are involved (multi-trait, multi-gene); we are using an integrated approach to test:

- Physiological parameters
- Genetical inheritance and interaction between loci
- Gene expression (genomics)

Using the power of Physiological Genetical Genomics
Plant Material

- RIL population of *Solanum lycopersicum* – *Solanum pimpinellifolium*
- 101 lines – F8
- Genotyping in F7 with 417 markers (mainly AFLP markers)
Physiological parameters

**Germination (potential)**
- t1 (time to germination of 1st seed)
- t50 (mean time of germination)
- u7525 (uniformity of germination)
- Gmax (maximum germination)
Dormancy (after-ripened vs. fresh seeds)

**Seed longevity**
Germination after controlled deterioration

**Reserve food**
- Seed size and seed mass

**Seedling morphology**
- Normal seedlings
- Cotyledon size
- Root growth rate
- Usable plants

**Related to vigour**
- Germination under osmotic stress
- Germination under salt stress
- Germination under temperature stress

**Assimilate partitioning (metabolomics)**
- Mono-, di- and oligosaccharide content
- Total lipid content
- Total protein content

**Key enzymes**
- Invertase
- Sucrose synthase
- ADP-glucose pyrophosphorylase (AGPase)
QTLs found

<table>
<thead>
<tr>
<th>Condition</th>
<th>#QTLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal conditions</td>
<td>6</td>
</tr>
<tr>
<td>t50</td>
<td>6</td>
</tr>
<tr>
<td>u7525</td>
<td>5</td>
</tr>
<tr>
<td>High temperature</td>
<td>4</td>
</tr>
<tr>
<td>Gmax</td>
<td>4</td>
</tr>
<tr>
<td>t50</td>
<td>2</td>
</tr>
<tr>
<td>u7525</td>
<td>3</td>
</tr>
<tr>
<td>Low temperature</td>
<td>1</td>
</tr>
<tr>
<td>Gmax</td>
<td>1</td>
</tr>
<tr>
<td>t50</td>
<td>5</td>
</tr>
<tr>
<td>u7525</td>
<td>1</td>
</tr>
<tr>
<td>Osmotic stress</td>
<td>3</td>
</tr>
<tr>
<td>Gmax</td>
<td>3</td>
</tr>
<tr>
<td>t50</td>
<td>3</td>
</tr>
<tr>
<td>Salt stress</td>
<td>6</td>
</tr>
<tr>
<td>Gmax</td>
<td>6</td>
</tr>
<tr>
<td>t50</td>
<td>4</td>
</tr>
<tr>
<td>u7525</td>
<td>3</td>
</tr>
<tr>
<td>Seed weight</td>
<td>4</td>
</tr>
</tbody>
</table>

Explained variances:
varying from 12 to 78 %
Confirmation and fine mapping

Inbred Backcross Lines (IBL)
- *Solanum lycopersicum – Solanum pimpinellifolium*
  - Tanksley population: 91 lines
  - Mapped with 127 markers (mainly CAPS markers)

Heterogeneous inbred families : HIFs
- Use residual heterozygosity for verification and fine mapping
- ~50% of tomato genome covered
- Genome wide fixed set is available in Arabidopsis

Verification of phenotype and test for interacting loci

Confirmation and fine mapping: HIF lines

QTL chrom 2

RIL

Line #33

HIF

Use recombination in this region for fine mapping

Screen ~500 plants for 1mb mapping
Screen ~6000 plants for gene mapping

Verification of phenotype and test for interacting loci
Fine mapping does not allow a high-throughput analysis and does not provide an global overview of the molecular mechanisms involved.
Genomics

eQTL-study for tomato

Microarrays for tomato:

- **Tom1 array:** cDNA array 13K probes, ~8.5K genes
- **Tom2 array:** oligo array 11K probes, ~10K genes
- **Affymetrix genechip:** 10K probes, ~9.2K genes
- **Cobimatrix:** oligo array 90K probes, ~22K genes
- **New affymetrix available ????:** 30K-40K probes
Pepper Tiling Array GeneChip

- 6,473,556 features of genomic probes 25 bp in a 2 bp overlap tiling pattern
- Corresponding to 30,815 pepper transcripts
- ~50% of the probes hybridize to *S. lycopersicum* and *S. pimpinellifolium*
- 30,790 transcripts in tomato with an average of 70 probes per unigene

[Diagram showing probe data for different genomic elements: Capsicum ESTs 123,489, Capsicum genomic DNA 466, Capsicum mRNAs 515, COS markers 642.]

Probes synthesized on GeneChip®
Conclusions

- RIL lines of *Solanum lycopersicum* – *Solanum pimpinellifolium* show phenotypical differences for seed quality parameters
- These differences in combination with the size and resolution of the population are enough to find QTLs for seed quality traits
- Preliminary analysis of IL lines show similar and different QTLs for seed quality traits
- Availability of IBLs of *S. pimpinellifolium* in *S. lycopersicum* background and HIFs will facilitate faster fine-mapping
Future Plans

- Continue to phenotype RIL population in more detail
- Genotype RIL population via eQTL study
- Start fine-mapping of most interesting QTLs with help of HIFs and/or IBL lines
- Microarray analysis, “likely candidate gene approach” and synteny with *Arabidopsis* to help in finding the corresponding genes
Acknowledgements

STW project members:
Natural variation in optima forma

Promoters
Linus van der Plas – WUR (RJ)
Harro Bouwmeester – WUR (RK, NK)

Advise
Maarten Koornneef - MPI Cologne
Joost Keurentjes – WUR

Statistics
Ritsert Jansen – RUG

Microarray analysis
Nicholas Provart – Toronto

Mapping populations
Sjaak van Heusden – PRI
Olivier Loudet - INRA