Phenotypes and genebanks

the importance and challenges of phenotyping Plant Genetic Resources

Theo van Hintum

20 April 2016

PhenomenAll COST Workshop, Copenhagen

WAGENINGEN **UR**
For quality of life

Centre for Genetic Resources, the Netherlands

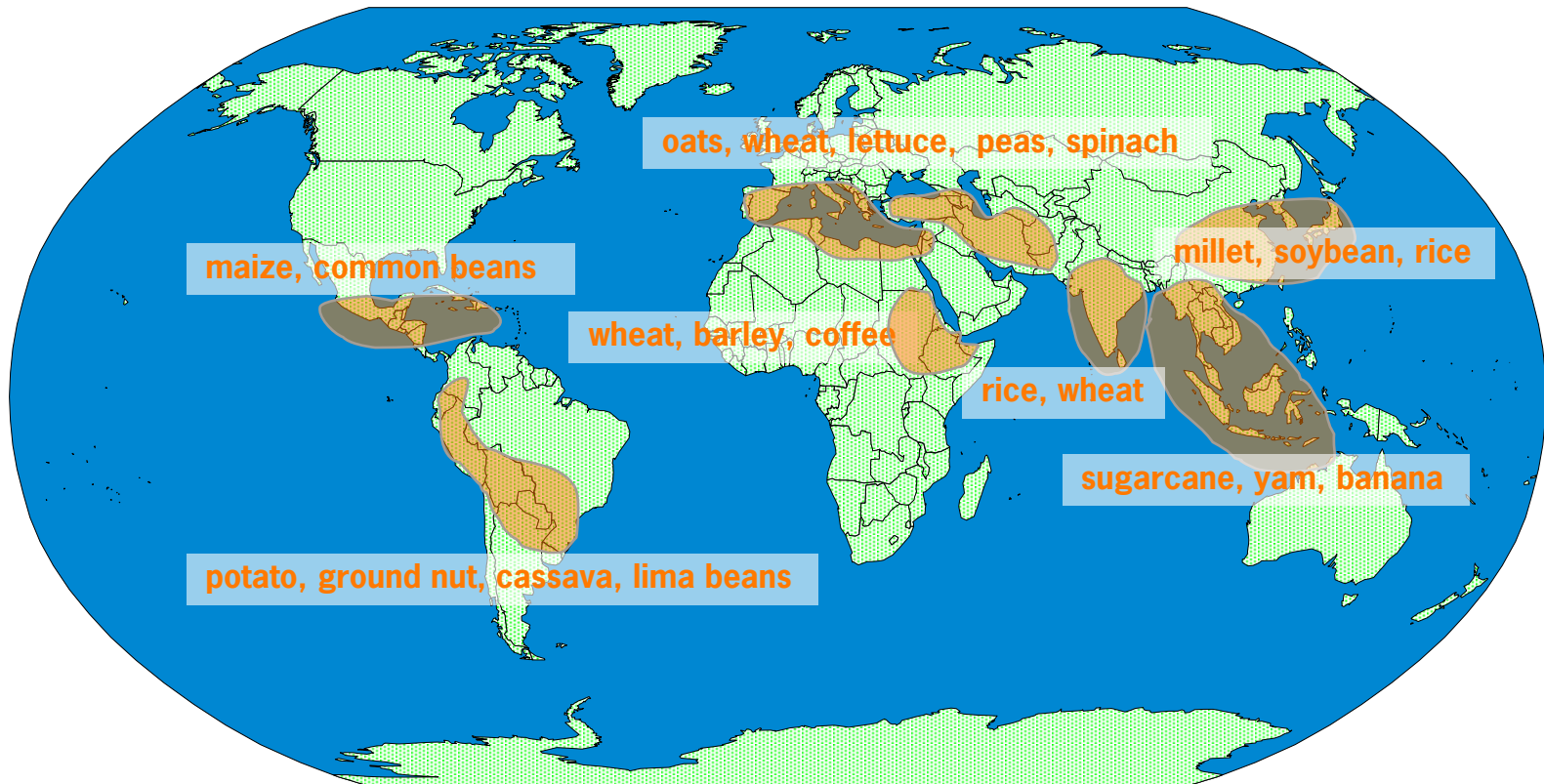# Phenotypes and genebanks

- **background**
  - plant genetic resources conservation
    - why ?
    - global effort
    - genebanks' role

- **phenotypes in genebanks**
  - current situation
    - type of data and availability
  - future scenario's
    - linking to genomics community

WAGENINGEN UR
For quality of life

Phenotypes and genebanks

- crop plant genetic diversity in historical context
  - farmers have been domesticating crops since c. 10,000 years
    - selection adapted phenotypes from wild populations
    - in 'centers of origin'
    - result: landraces

Phenotypes and genebanks

- *Vaviliov's centres of origin*

oats, wheat, lettuce, peas, spinach

millet, soybean, rice

maize, common beans

wheat, barley, coffee

rice, wheat

sugarcane, yam, banana

potato, ground nut, cassava, lima beans

WAGENINGEN UR
For quality of life

Phenotypes and genebanks

- crop plant genetic diversity in historical context
  - since c. 1850 crop genetic diversity disappeared due to monocultures and habitat destruction
    - scientific plant breeding based on Mendel and de Vries
    - urbanization, desertification, climate change
    - result: genetic erosion

Phenotypes and genebanks

■ genetic diversity needs to be conserved and made available

- *ex situ* – in genebanks

- *in situ* – in nature reserves

- *on-farm* – in 'traditional' farming systems

WAGENINGEN UR
For quality of life

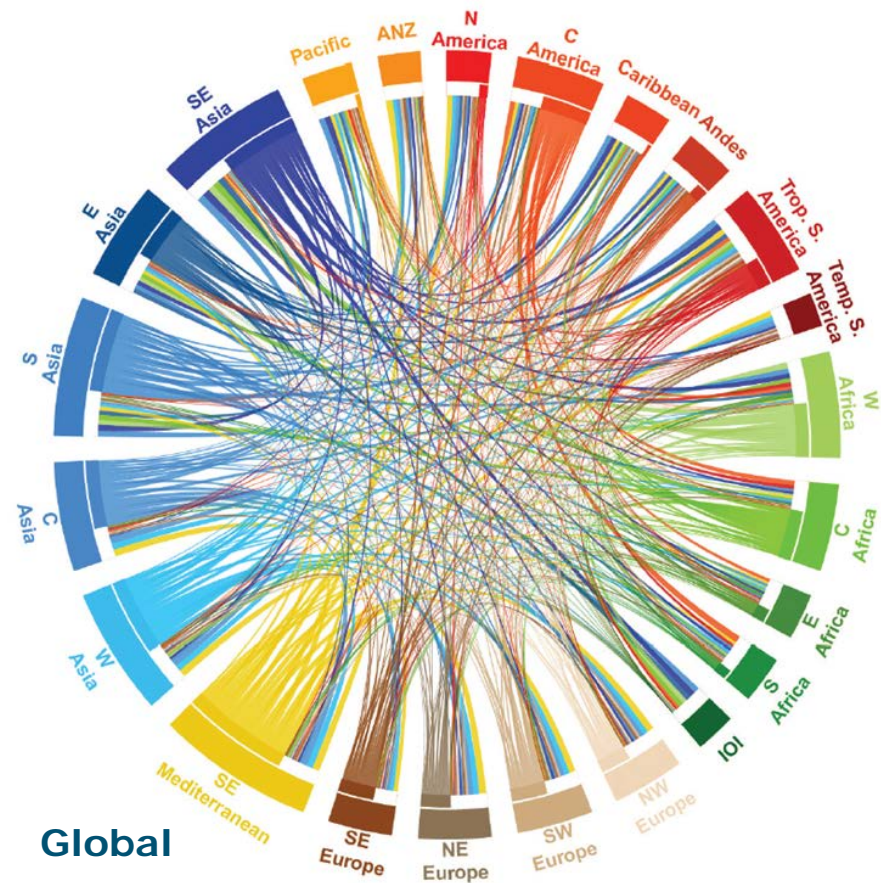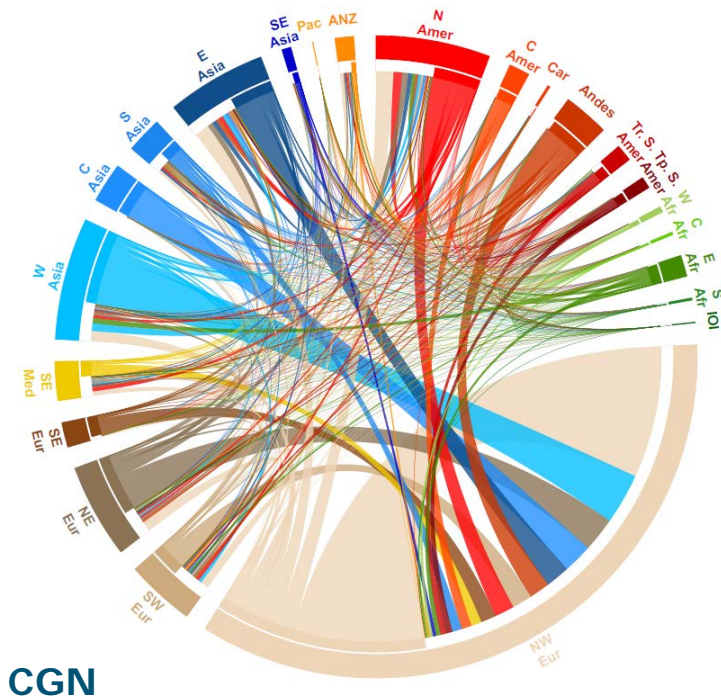# Phenotypes and genebanks

- *ex situ* conservation of crops: a global effort
  - world-wide network of genebanks (>10k accs)
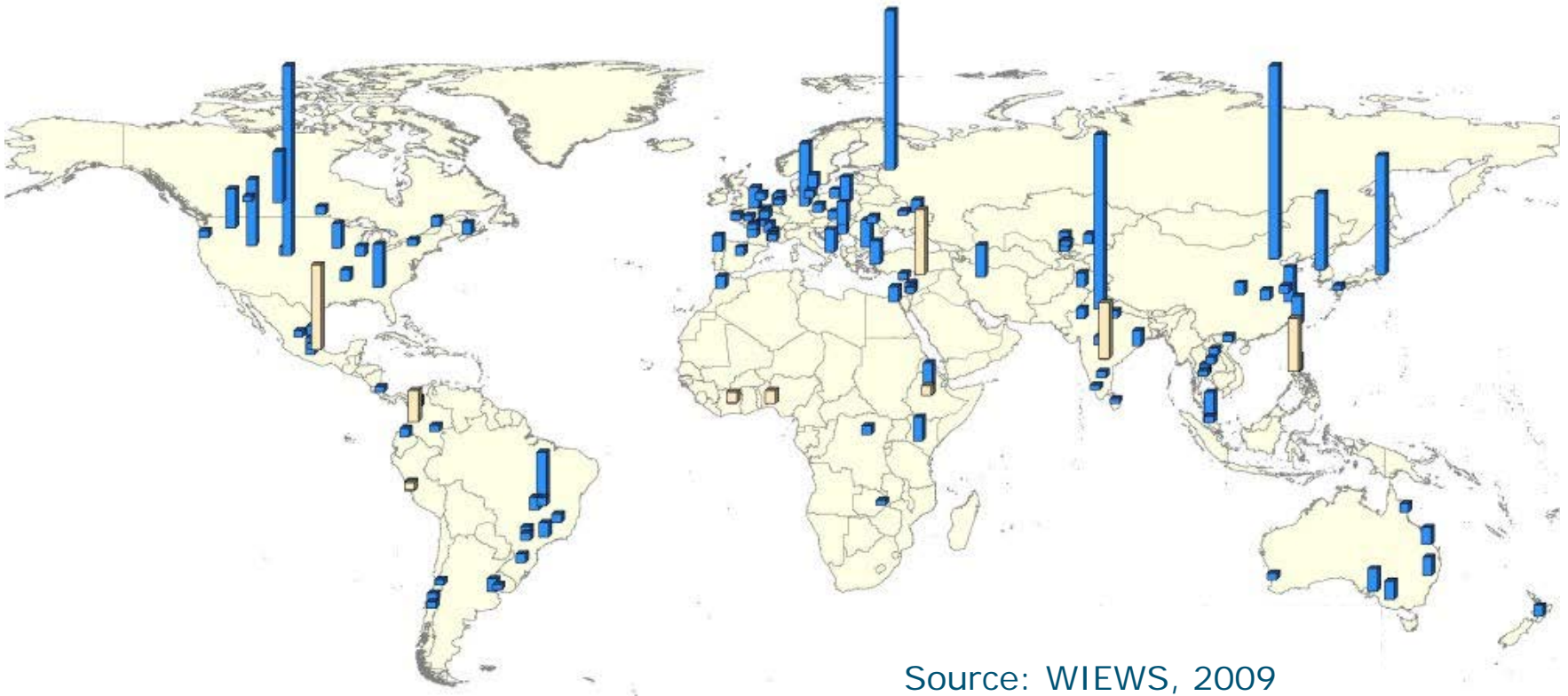
Source: WIEWS, 2009

Phenotypes and genebanks

- plant genetic resources
  - global interdependence

CGN

Global

Source: Khoury et al. 2015

Phenotypes and genebanks

- *ex situ* conservation of crops: a global effort
  - permafrost back-up facility at Svalbard

WAGENINGEN UR
For quality of life

Phenotypes and genebanks

- *ex situ* conservation of crops: a global effort
  - international legal framework (under developement)
    - Convention on Biological Diversity (CBD)
      - Nagoya Protocol Access and Benefit-sharing
    - International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA)

Phenotypes and genebanks

- genebank tasks: regeneration

self-fertilizing crop: barley

cross–fertilizing crops: potato, clover

WAGENINGEN UR
For quality of life

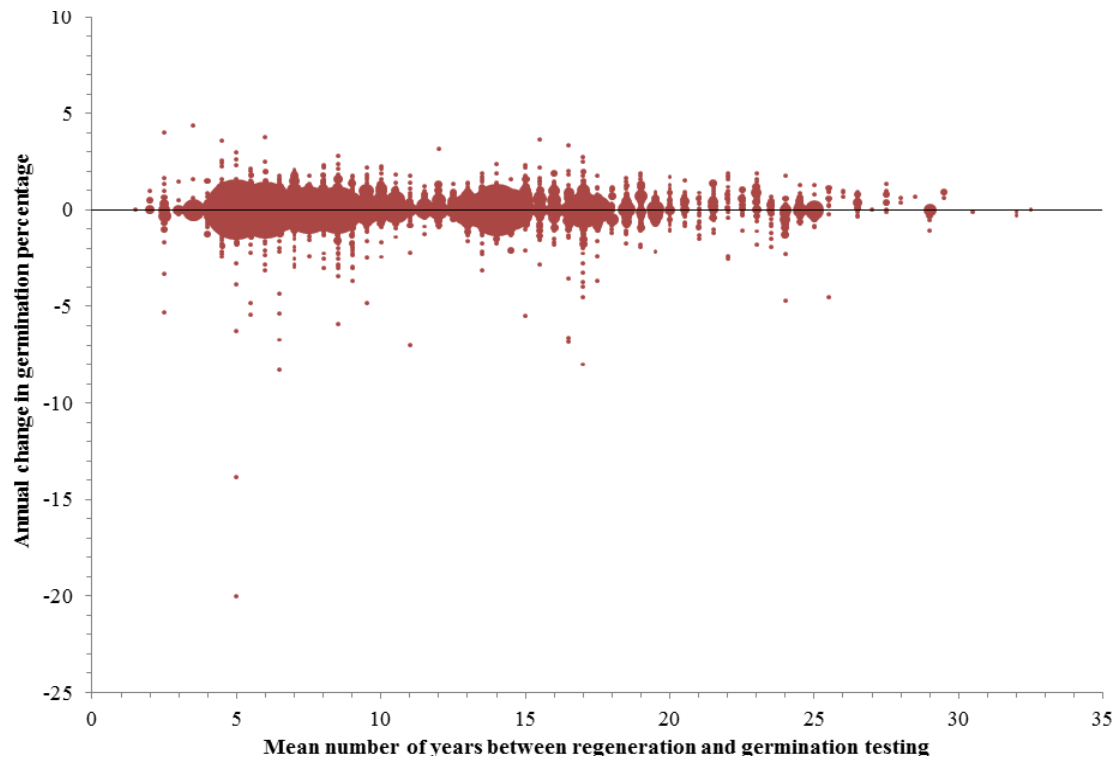# Phenotypes and genebanks

- genebank tasks: storing

packing in sealed bags

long term storage -20⁰C

Phenotypes and genebanks

■ germination testing

- no apparent change in germination due to storage

Phenotypes and genebanks

■ CGN: size collections (April 2016)

● lettuce	2401		● wheat	4908
● potato	1467		● barley	2666
● tomato	1330		● flax	952
● onion	417		● peas	1010
● spinach	468			
● sweet pepper	1032		● *other crops*	3442
● cucumber	926			
● cabbage etc.	1791			
			● TOTAL	22810

WAGENINGEN UR
For quality of life

Phenotypes and genebanks

- CGN: number of phenotypic datapoints per accession (April 2016)

WAGENINGEN UR
For quality of life
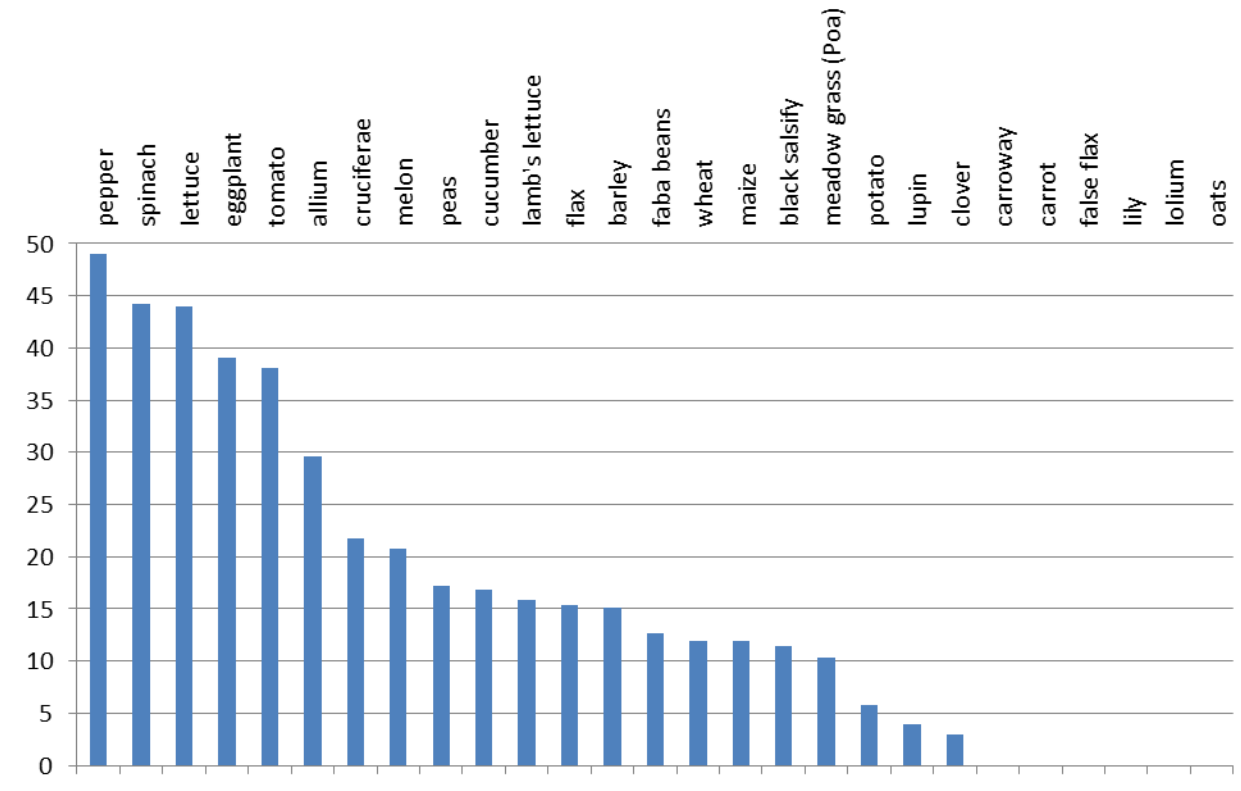
Phenotypes and genebanks

■ CGN: on-line access to data
  - all phenotypic data can be downloaded
  - summarized data (1-5 score) can be searched

  - <u>on-line demo</u> or <u>off-line demo</u>

Phenotypes and genebanks

- CGN: phenotypes versus distribution (April 2016)

Phenotypes and genebanks

- source of phenotypic data
  - in genebank jargon phenotypic data are C&E data: characterisation and evaluation data

  - characterisation
    - easy to score traits with high heritability
      - e.g. flower color, plant height
      - traits from descriptor lists (IPGRI, UPOV, etc.)
    - done by genebank to check identity and allow first selection

7.2.7   Leaf apex shape
(See Fig. 4)

      1   Acuminate
      2   Acute
      3   Apiculate
      4   Obtuse

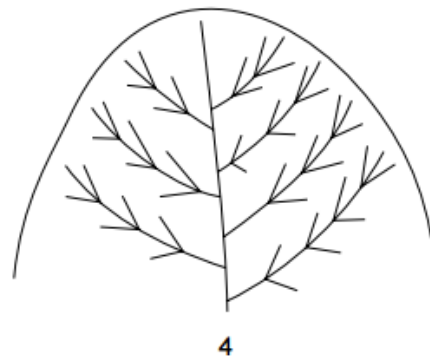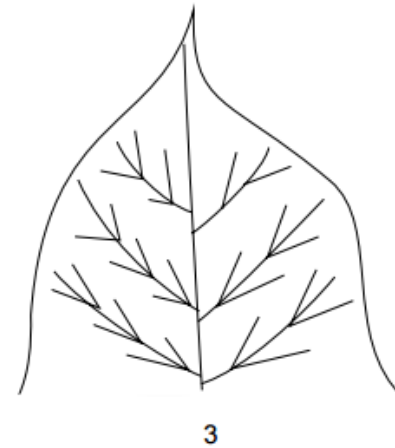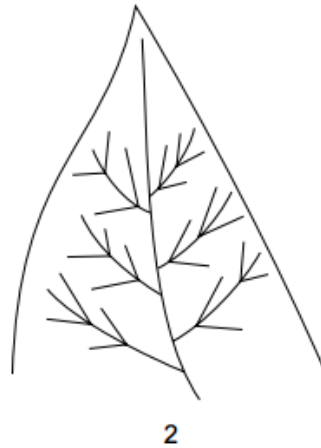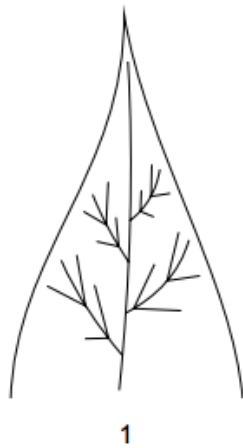Fig. 4. Leaf apex shape

Phenotypes and genebanks

- source of phenotypic data
  - in genebank jargon phenotypic data are C&E data: characterisation and evaluation data

  - evaluation
    - traits relevant to use
      - e.g. disease resistance, yield
    - done by crop experts in collaborative projects (PPPs, EU projects, etc.) or as feedback under the SMTA

WAGENINGEN UR
For quality of life

Phenotypes and genebanks

- genebank tasks: characterisation and evaluation

characterisation: lettuce
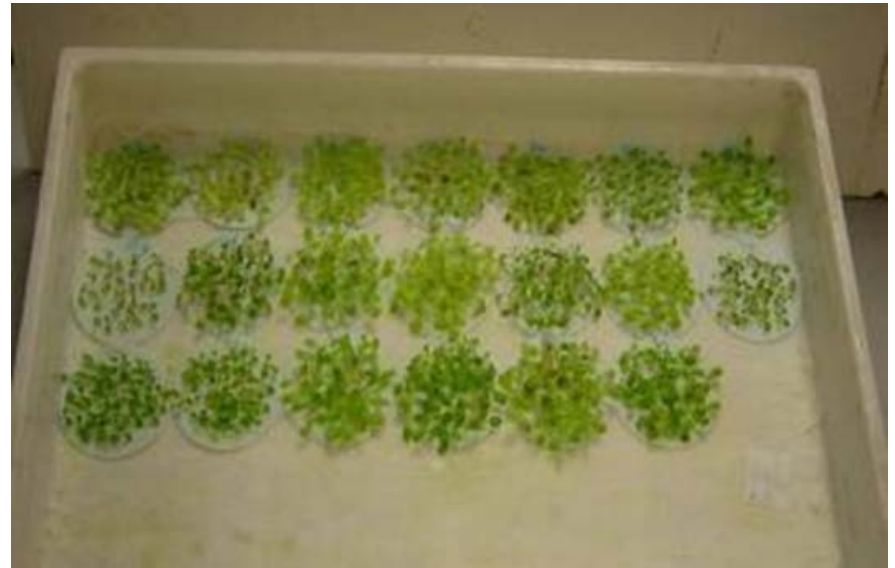
evaluation in the lab: lettuce

WAGENINGEN UR
For quality of life

Phenotypes and genebanks

- collaboration Dutch Breeders – CGN
  - collaboration is formalised in Letter of Intent/Memorandum of Understanding
    - currently 7 plant breeding companies
    - companies regenerate and do pre-competitive evaluations
      - under the umbrella of Plantum

WAGENINGEN UR
For quality of life

Education &
Programmes

Research &
Results

Expertise &
Services

Public-private partnerships with Dutch plant and animal breeding companies

Contact
dr.ir. L (Bert) Visser

Contact form

CGN has developed public-private partnerships with Dutch plant and animal breeding companies for the development, maintenance and evaluation of its collections.

✉ mailen   G+1   🐦 Tweet   in Share   f Like   0

Regeneration and improving the collection

Plant and animal breeding companies in the Netherlands contribute cash or in-kind to the build-up and management of the CGN collections in order to make the material available for research and breeding purposes.

Evaluation

Plant and animal breeding companies carry out evaluations on material incorporated in the CGN collections. The results of these evaluations may be published on the CGN website.

Collecting

Phenotypes and genebanks

- **issues**
  - low availability of data at data source
    - often in local poorly decoded Excel files – or worst
    - culture of lacking transparancy
    - making data accessible is not part of project
  - low reproducability (and thus useability) of large data sets
    - much material – bad experimental design – high GxE and σ_E
  - low combinability of datasets
    - local definitions of traits and methods and local GxE

Phenotypes and genebanks

- **issues**
  - difficulty in making many small data sets (with high relevance to user) and many large data sets (with low reliability and relevance) into a data set that can used to select material from large collection
    - e.g. 'lettuce leaf color' had 1402 accessions and 17 data sets but only 2082 observations (42 to 178 per dataset, 1 to 15 per accession) -> 91% missing values in 1402 x 17 matrix
    - Big Data / AI research might present solutions (statistics didn't sofar)

Phenotypes and genebanks

- **issues**
  - example dataset: 30 accessions, 10 experiments, 79 observations (73% missing values)
    - brightness = reliability
  - needed: algorithm to make consensus scores – highlighting conflicting results
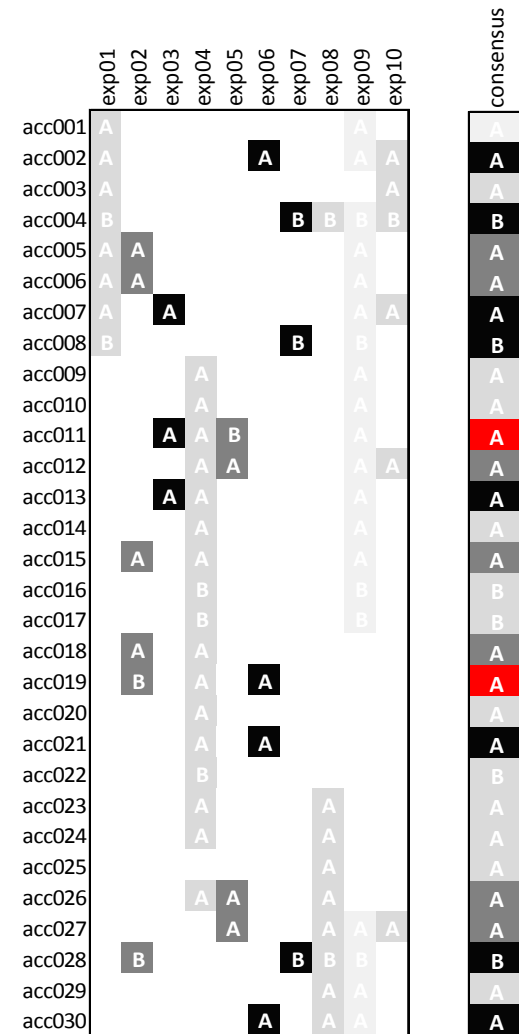
Phenotypes and genebanks

- interaction with genomics community
  - genomics needs genetic diversity and phenotypes
    - diversity from genebanks
    - phenotypic data is usually missing
      - no public databases - there's no phenotype GenBank
  - genebanks need genomics community to predict phenotypes
    - genebank collections are being (re-)sequenced

Phenotypes and genebanks

- **interaction with genomics community**
  - initiatives under way to make genebank data FAIR
    - findable, accessible, interoperable and re-useable
      - introduction Permanent Unique Identifiers
      - creation and application of ontologies
      - CGN: semantic annotation of all genebank webpages
    - inclusion of C&E data in EURISCO and GeneSys
  - DivSeek initative to bridge the software and standards gap
    - between genebank and genomics (& phenomics?)
    - basis: linked FAIR databases

WAGENINGEN UR
For quality of life

Div Seek

The future of food

RESIDES WITHIN CROP DIVERSITY

LEARN MORE

What is DivSeek?

DivSeek will help to bridge the gap between the information requirements of genebank curators, plant breeders and more targeted upstream biological researchers, to support applied germplasm curation, forward-looking breeding programs and strategic research.

...k the potential of crop ...ebanks around the ...ilable to all so that it ...ance the productivity ...ience of crops and

DivSeek will help to bridge the gap between the information requirements of genebank curators, plant breeders and more targeted upstream biological researchers, to support applied germplasm curation, forward-looking breeding programs and strategic research.

"

DivSeek will focus a powerful beam of light into the depths of gene pools essential for human survival.

The DivSeek initiative will work with existing, emerging and future initiatives to characterize crop diversity and develop a unified, coordinated and cohesive

Phenotypes and genebanks

- interaction with phenomics community
  - at this meeting: impressed by the tools being developed & enthousiasm
  - generally: very few interactions
    - none with impact
    - need for better phenotypes will rise with availability of genomic datasets
    - phenotypes ≠ terabytes of pictures

      reducing the data to useable information is major challenge

      alternative: make data FAIR and rely on Big Data analysis

Centre for Genetic Resources, the Netherlands

Phenotypes and genebanks

- conclusions

  - generating data points and pictures is not the issue, making them useful or FAIR is !

  - integration of phenotypic data sets with each other and with genomic data sets is essential

There are two basic rules for giving a presentation:

1 – don't tell everything

feel free to ask the rest !

Basket   Log in

All Journal Articles and Book Content ▼   [   ]   Search   Advanced Search

Plant Genetic Resources
Characterization and Utilization

Home > Plant Genetic Resources

Welcome to Cambridge Journals Online

To access subscriptions and personalised features please log in or register

Register for an Account

f y M ✉ +

Like 13   G+1 2

JOURNAL INFORMATION

JOURNAL MENU

Plant Genetic Resources Home
Save Journal to My CJO Account

Browse Journal
FirstView Articles
Latest Issue
Back Issues

Browse Articles

Plant Genetic Resources
Characterization and Utilization

New Content Alerts

Editor(s):   Theo van Hintum, *Centre for Genetic Resources, Wageningen University and Research Centre, Netherlands*
Editorial Board

Published on behalf of the National Institute of Agricultural Botany (NIAB)

Plant Genetic Resources is an international journal which provides a forum for describing the application of novel genomic technologies, as well as their integration with established techniques, towards the understanding of the genetic variation captured in both in situ and ex situ collections of crop and non-crop plants; and for the airing of wider issues relevant to plant germplasm conservation and utilisation. We particularly welcome multi-disciplinary approaches that incorporate both a technical and a socio-economic focus. Technical aspects can cover developments in technologies of potential or demonstrated relevance to the analysis of variation and diversity at the phenotypic and genotypic levels; the development of rational germplasm collection, evaluation and conservation strategies; and
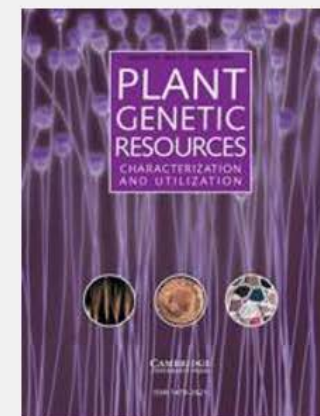
ISSN: 1479-2621   EISSN: 1479-263X
Frequency: 3 Issue(s) per year

Open Access options available   READ the Instructions for Authors   Register for FREE Content Alerts

Thank you for your attention !

Genetic resources are useful, pretty and tasteful

Genetic resources are the food on your plate

Thank you for your attention !