

European genebanks in the genomics era

Theo van Hintum

Centre for Genetic Resources, The Netherlands (CGN)



European genebanks in the genomics era

- the world faces 'global grand challenges'
 - feed the world in a sustainable way
 - population size
 - climate change
 - many potential solutions
 - explore the potential of diversity in genebanks
 - first focus on CGIAR genebanks (714k accessions)
 - other genebanks will also get involved (~6mio accessions)
 - use new technology: genotyping, phenotyping, ICT

Susan McCouch: *"Sequencing will focus a powerful beam of light into the depths of gene pools essential for human survival"*

European genebanks in the genomics era

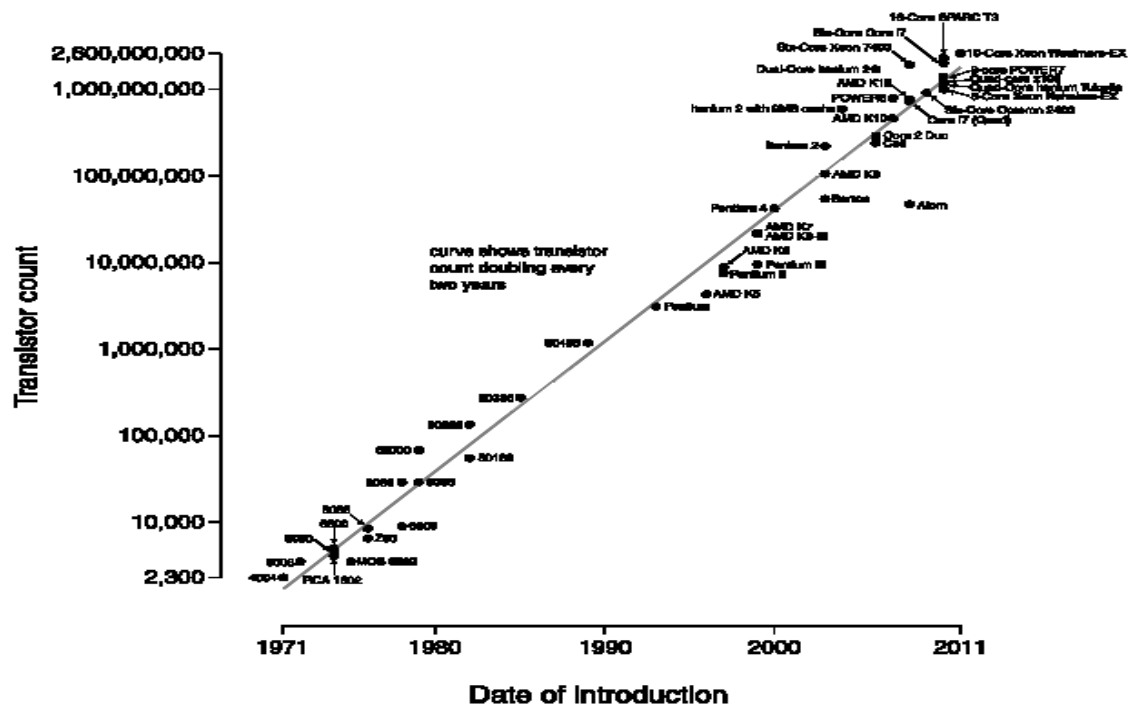
- plant breeding is changing / has changed
 - classical breeding
 - based on phenotypic information
 - $A \times B$, some selfing and select the best
 - $A^n \times B$ and keep the desired trait from B
 - modern breeding
 - based on phenotypic and sequence information
 - marker assisted breeding: $A^n \times B$ and keep the desired trait from B by following it with genetic markers
 - trait identification via GWAS, allele mining via genomic identification of functional SNPs, comparative genomics
 - genomic selection by defining the breeding value as a function of the entire genome

European genebanks in the genomics era

■ Moore's law

- the number of transistors on integrated circuits doubles approximately every two years

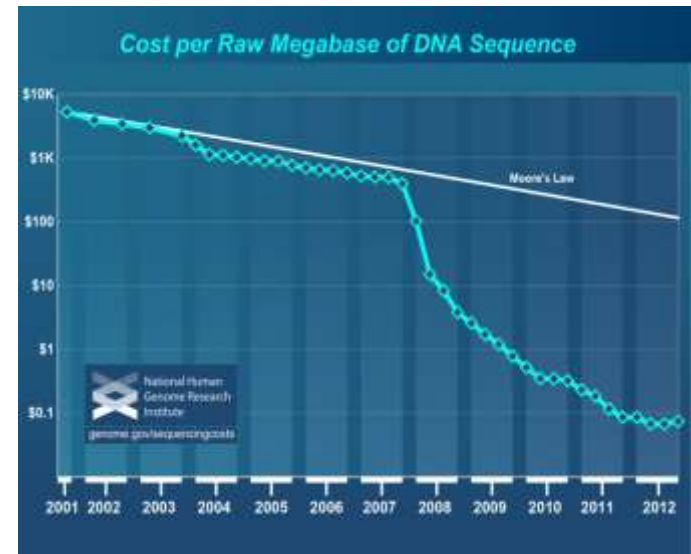
Microprocessor Transistor Counts 1971-2011 & Moore's Law



European genebanks in the genomics era

■ Moore's law

- sequencing goes even faster !
- the research community will be / is flooded with sequencing data
 - computers will have difficulty to cope
- if we want to benefit, we'd better be prepared



European genebanks in the genomics era

■ genebanks will change



The International Center for Tropical Agriculture in Colombia holds 65,000 crop samples from 141 countries.

Feeding the future

We must mine the biodiversity in seed banks to help to overcome food shortages, urge **Susan McCouch** and colleagues.

Humanity depends on fewer than a dozen of the approximately 300,000 species of flowering plants for 80% of its caloric intake. And we capitalize on only a fraction of the genetic diversity that resides within each of these species. This is not enough to support our food system in the future. Food availability must double in the next 25 years to keep pace with population and income growth around the world. Already, food-production systems are precarious in the face of intensifying demand, climate change, soil degradation and water and land shortages.

Farmers have saved the seeds of hundreds of crop species and hundreds of thousands of 'primitive' varieties (local domesticates called

landraces), as well as the wild relatives of crop species and modern varieties no longer in use. These are stored in more than 1,700 gene banks worldwide. Maintaining the 11 international gene-bank collections alone costs about US\$18 million a year.

The biodiversity stored in gene banks fuels advances in plant breeding, generates billions of dollars in profits, and saves many lives. For example, crossbreeding a single wild species of rice, *Oryza nivara*, which was found after screening more than 6,000 seed-bank accessions, has provided protection against grassy stunt virus disease in almost all tropical rice varieties in Asia for the past 36 years¹. During the green revolution, high-yielding rice and wheat varieties turned India into a net

food exporter. By 1997, the world economy had accrued annual benefits of approximately \$115 billion from the use of crop wild relatives² as sources of environmental resilience and resistance to pests and diseases.

The time is ripe for an effort to harness the full power of biodiversity to feed the world. Plant scientists must efficiently and systematically domesticate new crops and increase the productivity and sustainability of current crop-production systems.

Why does plant breeding need a boost? Because new, high-yielding seeds that are adapted for future conditions are a cornerstone of sustainable, intensified food production³. Since the mid-1990s, progress in conventional plant breeding has ▶

4 JULY 2013 | VOL 499 | NATURE | 23

European genebanks in the genomics era

- genebanks will change

© NIAB 2014
ISSN 1479-2621

Plant Genetic Resources: Characterization and Utilization; 1–10
doi:10.1017/S1479262114000082

Next-generation genebanking: plant genetic resources management and utilization in the sequencing era

Rob van Treuren* and Theo J. L. van Hintum

Centre for Genetic Resources, the Netherlands (CGN), Wageningen University and Research Centre, PO Box 16, NL-6700 AA Wageningen, The Netherlands

Received 6 November 2013; Accepted 1 February 2014

Abstract

Advances in sequencing technologies have made it possible to analyse large amounts of germplasm against low production costs, which has opened the door to screen genebank collections more efficiently for DNA sequence variation. The present study explores how these developments may affect the operations of genebanks and, consequently, how genebank agendas may change. It is argued that the new developments are more likely to have

European genebanks in the genomics era

- potential impact genomics
 - bad scenario (current situation?)
 - genomics community gets material from genebanks
 - lines are selected, populations are created
 - this is genotyped, phenotyped and analysed
 - material is maintained (and lost) by the research community
 - one way flow from genebanks to research community
 - no synergistic effects

Genebanks & genomics

- potential impact genomics
 - good scenario
 - genebanks and genomics community collaborate in selecting and preparing material for research
 - high quality (authentic, viable)
 - homozygous and homogeneous (if appropriate & possible)
 - research populations
 - genomics community does research
 - genebanks taps into results
 - better serve users – including future genomic research

European genebanks in the genomics era

■ the challenges

- unlock the genetic diversity conserved in genebanks for use by breeders via genomic science
- make sure the data / knowledge generated improves the availability of germplasm

■ issues to be addressed

- access to germplasm
- data generation
- data sharing and use

European genebanks in the genomics era

- access to germplasm
 - genebanks can prepare their germplasm to make it 'genomic analysis ready'
 - assure access to collection (information and material)
 - select homogeneous lines from accessions
 - verify authenticity of accessions
 - optimise represented diversity
 - make research populations available
 - issues
 - what do genebanks need to know & how can they get that knowledge?
 - how will these activities be organised and funded?

European genebanks in the genomics era

■ data generation

- the genebank material needs to be characterised
 - genotypic and phenotypic
 - with proper quality assurance
- issue
 - who will initiate, do and fund the data generation?

European genebanks in the genomics era

- data sharing and use
 - the resulting data (and the subsequent analysis) need to be made accessible
 - platforms and databases
 - standards
 - IP solutions
 - methods for analysis need to be developed and made available
 - issue
 - how can this be organised, by whom, with what funds?

European genebanks in the genomics era

- all the issues were recognised by a group of scientists
 - SeedSeq and the Digital Seed Bank initiatives
 - both wanted to bridge the gap between PGR and genomic community
 - DivSeek initiative
 - merges SeedSeq and the Digital Seed Bank

European genebanks in the genomics era

■ DivSeek

- mission

“to enable plant breeders and researchers to mobilize a vast range of plant genetic variation to accelerate the rate of crop improvement and furnish food and agricultural products to the growing human population.”

- strategy

1. managing and preparing genebank collections effectively to promote their use and to facilitate comprehensive genotypic and phenotypic analyzes
2. performing a detailed characterisation of the diversity that exists in gene banks so that it can be targeted for use in crop breeding programs

European genebanks in the genomics era

■ DivSeek

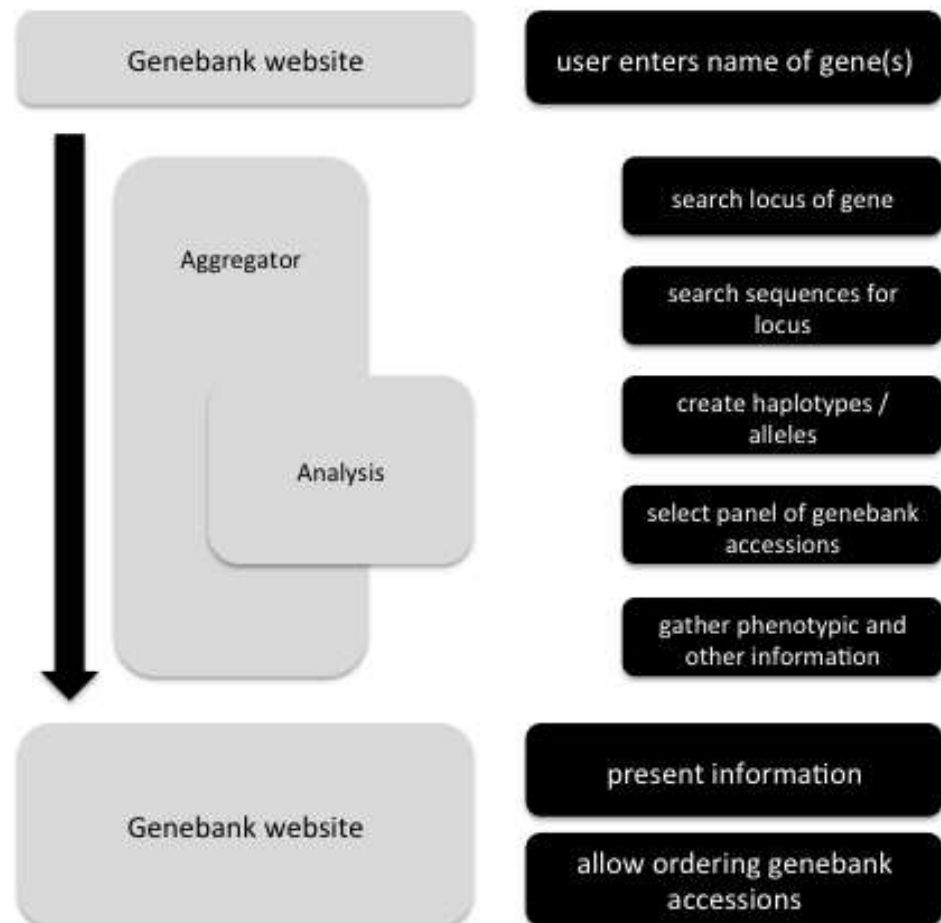
- status
 - two meetings were held
 - January 2014, San Diego – 85 participants
 - April 2014, Bonn – 25 participants
 - White Paper is being finalized
 - program is being formulated and started
- coordination
 - first year: Global Crop diversity Trust
- funding
 - concept & programme development: GCDT
 - programme: yet unclear

European genebanks in the genomics era

- example approach from genebank environment
 - use case
 - genebank website allows the user to enter name of gene
 - interface returns list of accessions with different alleles for this gene
 - allele name, known function, information about the phenotype
 - R. Finkers, P.-Y. Chibon, R. van Treuren, R. Visser and T. van Hintum (*in press*) Genebanks and genomics: how to interconnect data from both communities? *Plant Genetic Resources*

European genebanks in the genomics era

- example approach from genebank environment



European genebanks in the genomics era

- example approach from genebank environment
 - all elements of this use case are available (for 150 tomato accessions)
 - material is available from CGN and can be ordered on-line
 - including SMTA, etc.
 - accession details are semantically annotated
 - currently only passport data
 - EU-SOL BreeDB has DNA sequences
 - prototype aggregator is available
 - presents CGN and EU-SOL data in one interface
 - analytical steps still need attention
 - haplotype creation \neq functional allele identification

European genebanks

- illustration of semantic annotation

CGN Accession details

Lycopersicon esculentum (Caro-Red, Intermediate Beta type; 327-b1)

Add

Order



This accession is not in the Shopping Cart

Identifiers

Accession number:	CGN22879
Accession name:	Caro-Red, Intermediate Beta type; 327-b1
Crop (and subcrop type):	tomato
Population type:	Research material

SMTA status

SMTA needed:	Yes
--------------	-----

Classification

Genus:	Lycopersicon
Species:	esculentum
Sub-taxon:	

Origin information

Collecting number:	
Country of origin:	
Collection site:	
Latitude / Longitude:	
Altitude (m):	
Collecting date:	
Donor institute:	Instituut voor de Veredeling van Tuinbouwgewassen, Wageningen, Netherlands
Donor number:	

European genebanks in the genomics era

- illustration of semantic annotation
 - view page source

```
view-source:applicaties.wageningenur.nl/applications/cgngenis/AccessionDetail.aspx?AccessionID=CGN22879

1
2 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN" >
3 <html version="HTML+RDFa 1.1" lang="en"
4   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
5   xmlns:pp="http://www.croponontology.org/rdf/CO_020:">
6   <HEAD>
7       <title>CGN Germplasm Orderservice - Details for
8           CGN22879
9       </title>
10      <meta name="GENERATOR" Content="Microsoft Visual Studio
11      <meta name="CODE_LANGUAGE" Content="C#">
12      <meta name="vs_defaultClientScript" content="JavaScript"
13      <meta name="vs_targetSchema" content="http://schemas.mic
14      <LINK href="Style/blue.css" type="text/css" rel="stylesh
15      </HEAD>
16      <body>
17          <form name="Form1" method="post" action="AccessionDetail.aspx?AccessionID=CGN22879"
18      <input type="hidden" name="__VIEWSTATE"
19      value="dDw5NDaxMDYyNDU7dDw7bDxpPDA+Oz47bDx0PDtsPGk8MT47aTwzPjtpPDU+O2k8O
20      z47aTwyOT47aTwyMT47aTwyNT47aTwyNz47aTwyOT47aTwyOMT47aTwyOMz47aTwyONT47aTwyON
21      iBkZXRhaWxzOz4+Oz47Oz47dDxwPHA8bDxUZXh0Oz47bDxcPGlcPkk5Y29wZXJzaWNvbiBlc
22      jctYjEpOz4+Oz47Oz47dDxwPHA8bDxJbWFfnZVVybDs+O2w8SW1hZ2VzL1Nob3BwaW5nY2Fyd
23      1Nob3BwaW5nY2FydC9TQ19Ob3Rjb1Nob3BwaW5nY2FydC5naWY7Pj47Pjs7Pjt0PHA8bDxUZ
24      WRpYXRlIEUldGEgdHlwZVw7IDMyNy1iMTs+Pjs7Pjt0PHA8bDxUZXh0Oz47bDx0b21hdG87P
25      Ds+O2w8WWVzOz4+Ozs+O3Q8cDxsPFRleHQ7PjtsPEX5Y29wZXJzaWNvbjs+Pjs7Pjt0PHA8b
26      DxsPFRleHQ7PjtsPFx1Oz4+Ozs+O3Q8cDxsPFRleHQ7PjtsPFx1Oz4+Ozs+O3Q8cDxsPFRle
27      Fx1Oz4+Ozs+O3Q8cDxsPFRleHQ7PjtsPFx1Oz4+Ozs+O3Q8cDxsPFRleHQ7PjtsPFx1Oz4+Ozs+O3Q8cDxsPFRle
28      H1dCB2b29yIGRlIFZlcmVkaW50b3R5b3BudWllZC19Mzc7Pj47Pjs7Pjs+Pjs+PjtsP
29      7PjtsPDEuMTEExNTs+Pjs7Pjt0PHA8b
30      8dD47Pj47Oz47dDxwPGw8VmlzaWJsZ
31      saXk1NTs+Pjs+Ozs+O3Q8cDxwPGw8V
32      Xl9Mzc7Pj47Pjs7Pjs+Pjs+PjtsP
33      />
34      <script language="javascript">
35      function openTraitInfo(url) {
```

European genebanks in the genomics era

- illustration of semantic annotation
 - view page source
 - scroll down

```
view-source:applicaties.wageningenur.nl/applications/cgngenis/AccessionDetails.aspx?ID=fzflly55&acnumber=CGN22879
<td class="content" style="WIDTH: 245px">SMTA needed:</td>
<td class="content">Yes</td>
</tr>
</table>
<br>
<table class="boxpart" width="600" border="0">
<tr>
<td class="kop" colspan="2">Classification</td>
</tr>
<tr>
<td class="content" style="WIDTH: 245px" valign="top">Genus:</td>
<td class="content" valign="top"><span property="pp:0000008">Lycopersicon</span></td>
</tr>
<tr>
<td class="content" style="WIDTH: 245px" valign="top">Species:</td>
<td class="content" valign="top"><span property="pp:0000006">esculentum</span></td>
</tr>
<tr>
<td class="content" style="WIDTH: 245px" valign="top">Sub-taxon:</td>
<td class="content" valign="top"><span property="pp:0000008"></span></td>
</tr>
</table>
<br>
<table class="boxpart" width="600" border="0">
<tr>
<td class="kop" colspan="2">Origin information</td>
</tr>
<tr>
<td class="content" style="WIDTH: 245px" valign="top">Collecting number:</td>
<td class="content" valign="top"></td>
</tr>
<tr>
<td class="content" style="WIDTH: 245px" valign="top">Country of origin:</td>
<td class="content" style="WIDTH: 245px" valign="top"></td>
</tr>
<tr>
<td class="content" style="WIDTH: 245px" valign="top">Genus:</td>
<td class="content" valign="top"><span property="pp:0000005">Lycopersicon</span></td>
</tr>
<tr>
<td class="content" style="WIDTH: 245px" valign="top">Species:</td>
<td class="content" valign="top"><span property="pp:0000006">esculentum</span></td>
</tr>
<tr>
<td class="content" style="WIDTH: 245px" valign="top">Sub-taxon:</td>
<td class="content" valign="top"><span property="pp:0000008"></span></td>
</tr>
</table>
```

European genebanks in the genomics era

■ conclusions

- genebanks can benefit from genomics revolution
 - to do their task better and stay relevant
- the genebank community should make sure it stays in dialogue with the genomics community
 - to avoid a one-way flow of material and information
- DivSeek is a promising initiative that will help the genebank community to become / stay connected

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 !

