

# Gap Analysis

Andy Jarvis, Julián Ramírez, Nora Castañeda, Robert Hijma, and Jacob Van Etten

# Why Gap Analysis?

- Tool to assess crop and crop wild relative genetic and geographical diversity
- Allows detecting incomplete species collections as well as defining **which** species should be collected and **where** these collections should be focused
- Assesses the **current extent** at which the *ex situ* conservation system is correctly **holding the genetic diversity** of a particular genepool

To know what you don't have, you first  
need to know what you do have

...and that's where things already get  
complex

# Welcome to database hell

News Feed

Status Updates

Photos

Posted Items

Live Feed



Andy Jarvis is about to talk about database hell. 11 minutes ago - [Comment](#)



Andrew Farrow is enthused by diffusion. about an hour ago - [Comment](#)

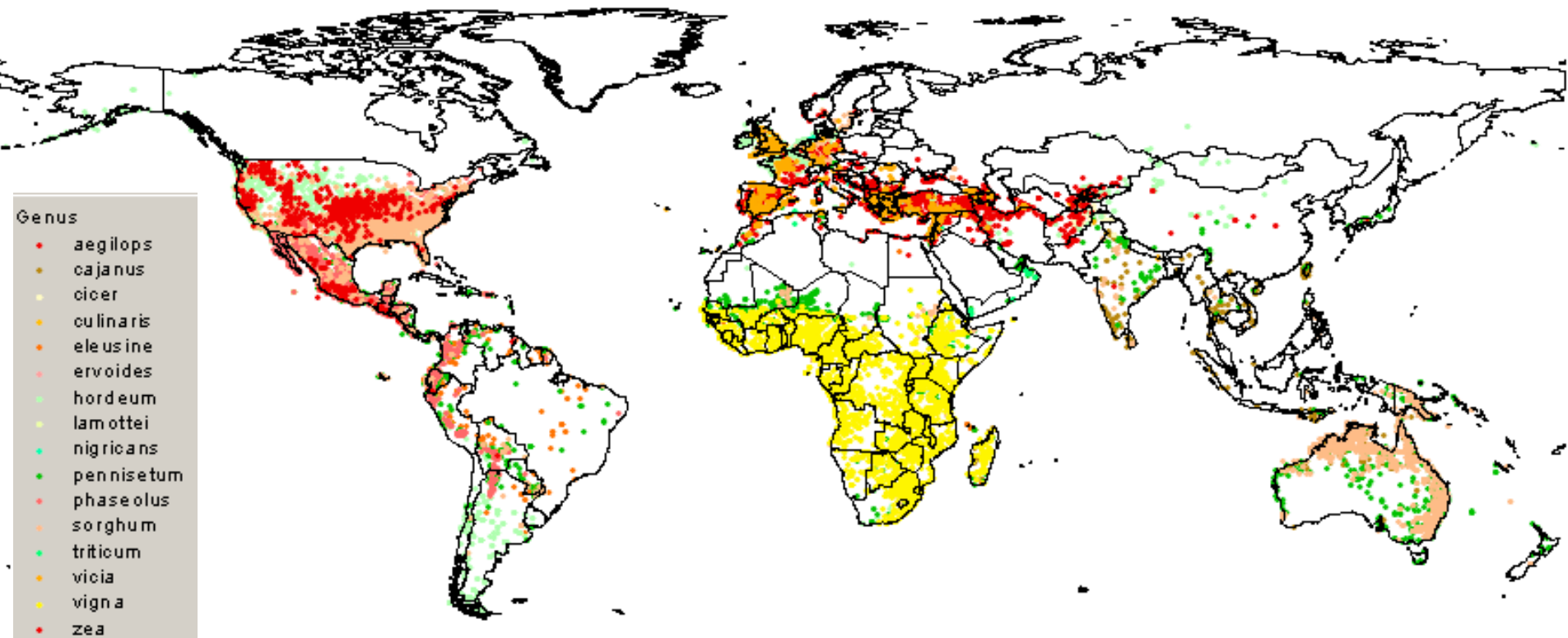


Luigi Guarino is in database hell. 4 hours ago - [1 Comment](#)

1. Árboles de la Península de Yucatán, Flora del Distrito de Tehuantepec, Oaxaca y Familia Asteraceae en México (IBUNAM)
2. Arizona State University Vascular Plant Herbarium
3. Biodiversidad de Costa Rica
4. California State University, Chico
5. CIAT-Genetic Resources Unit
6. Colecciones de George Boole Hinton depositadas en el herbario de Kew: Familia Leguminosae
7. Daniel Debouck - Cahiers de Phaséologie
8. DAO Herbarium Type Specimens
9. Ejemplares tipo de plantas vasculares del Herbario de la Escuela Nacional de Ciencias Biológicas, México (ENCB, IPN)
10. Fairchild Tropical Botanic Garden Virtual Herbarium Darwin Core format
11. [Freytag and Debouck \(2002\)](#)
12. GRIN
13. Harvard University Herbaria
14. Herbario de la Universidad de Arizona, EUA
15. Herbario del Instituto de Ecología, A.C., México (IE-BAJIO)
16. Herbarium (UNA)
17. Herbarium de Geo. B. Hinton, México
18. Herbier de la Guyane
19. Instituto de Ciencias Naturales
20. ITIS
21. Missouri Botanical Garden
22. National Botanic Garden of Belgium (NBGB)
23. National Vegetable Germplasm Bank, Mexico (BANGEV)
24. Native Seeds/SEARCH (NSS)
25. New York Botanical Garden (NYBG)
26. NMNH Botany Collections
27. Phanerogamie
28. Repatriación de datos del Herbario de Arizona (ARIZ)
29. Royal Botanic Gardens, Kew
30. The AAU Herbarium Database
31. The Deaver Herbarium, Northern Arizona University
32. United States National Plant Germplasm System Collection
33. USDA PLANTS Database
34. Vascular Plant Type Specimens

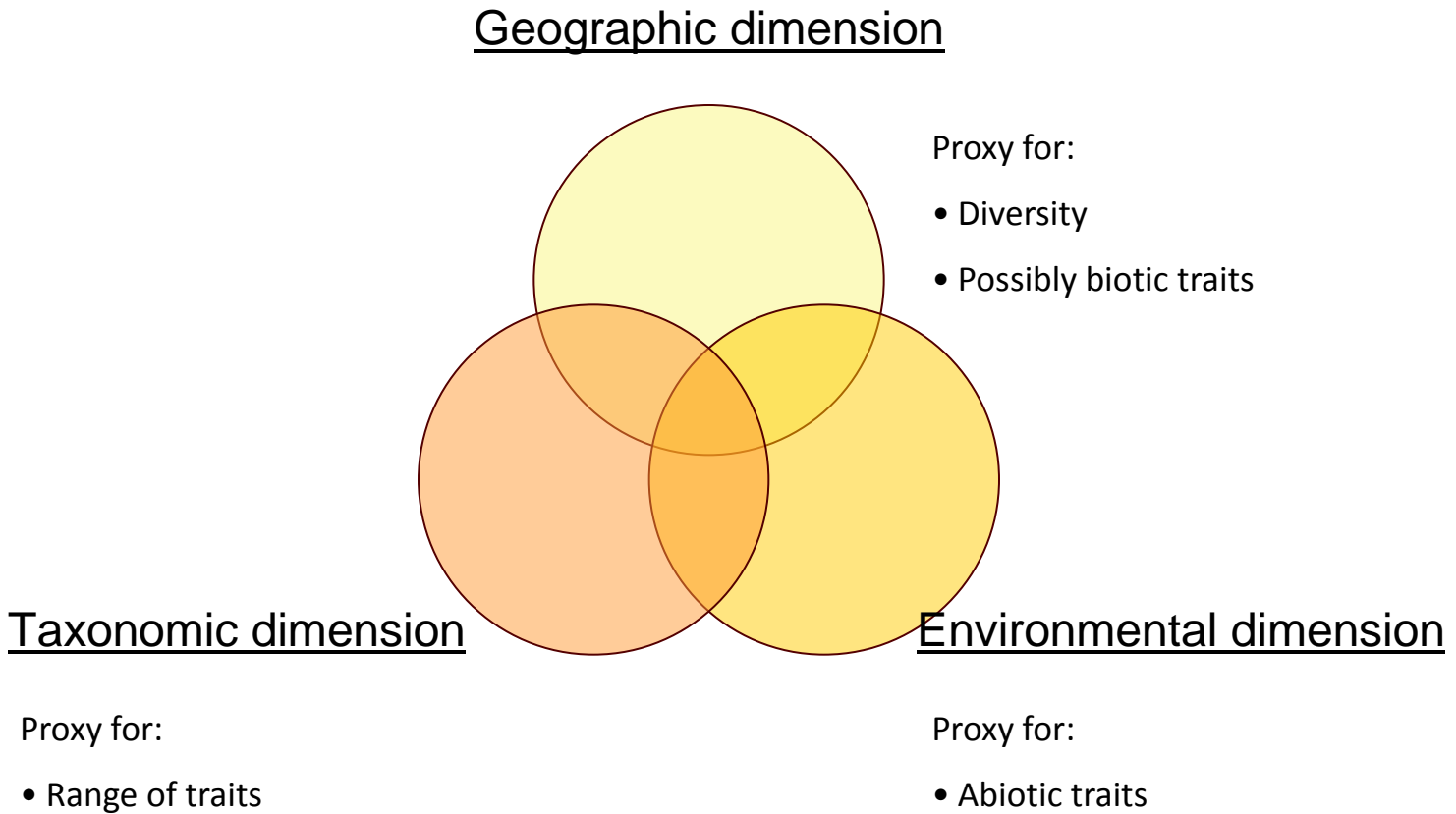


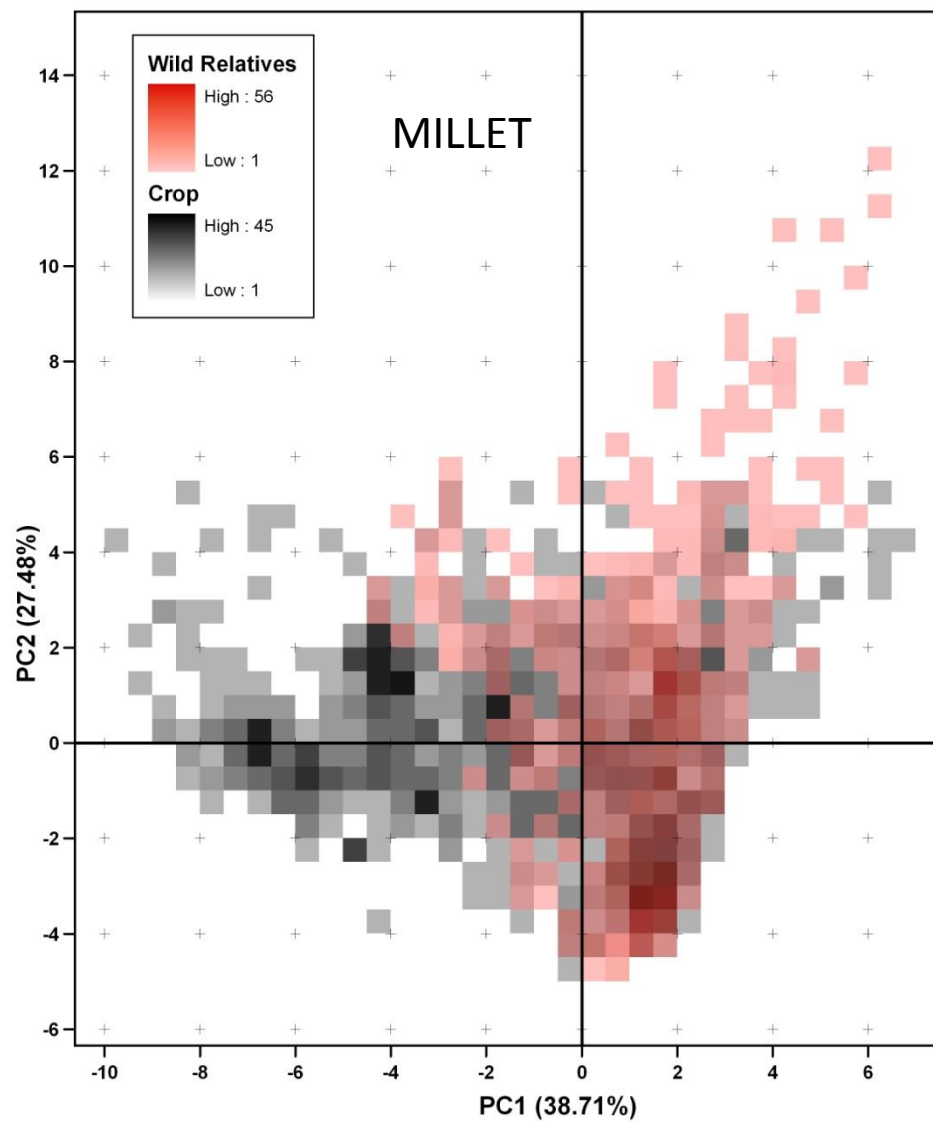
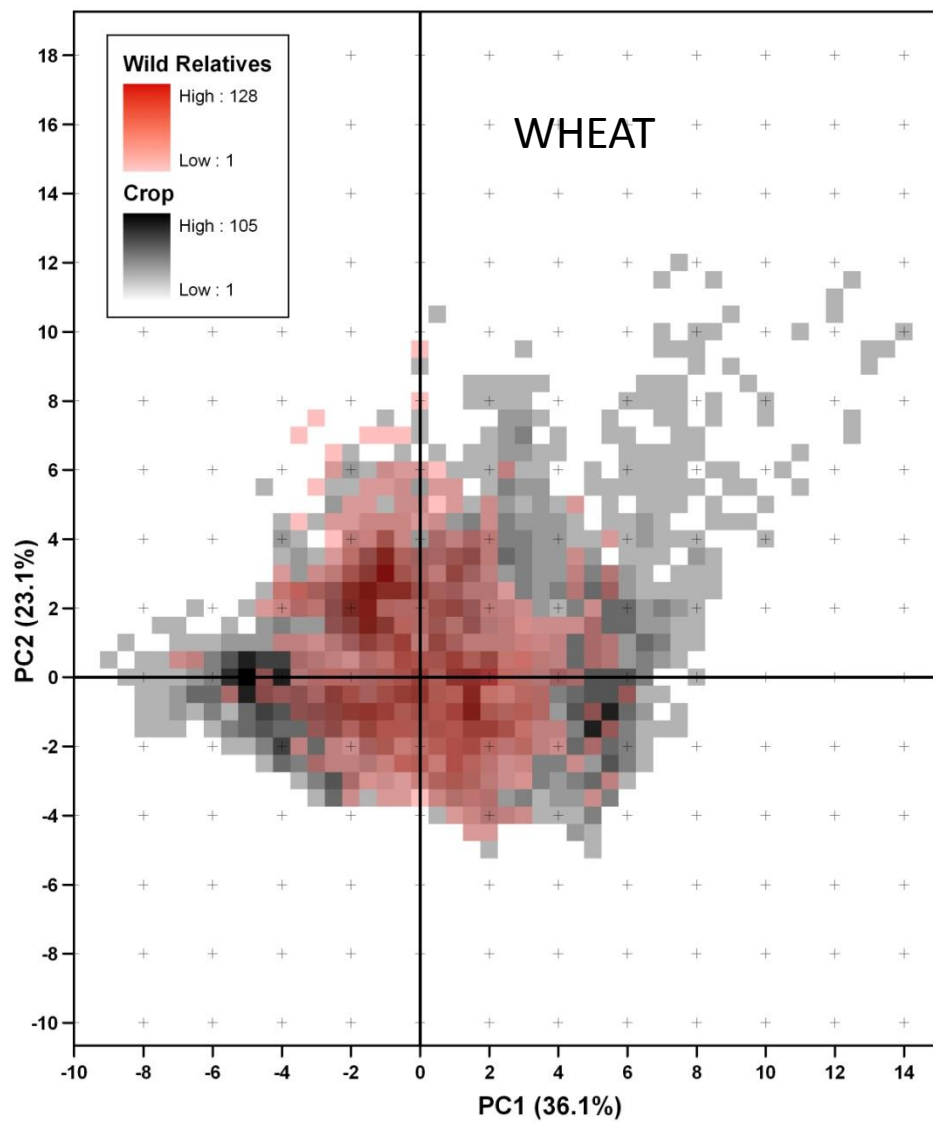
# The visible global system



# The Gap Analysis process

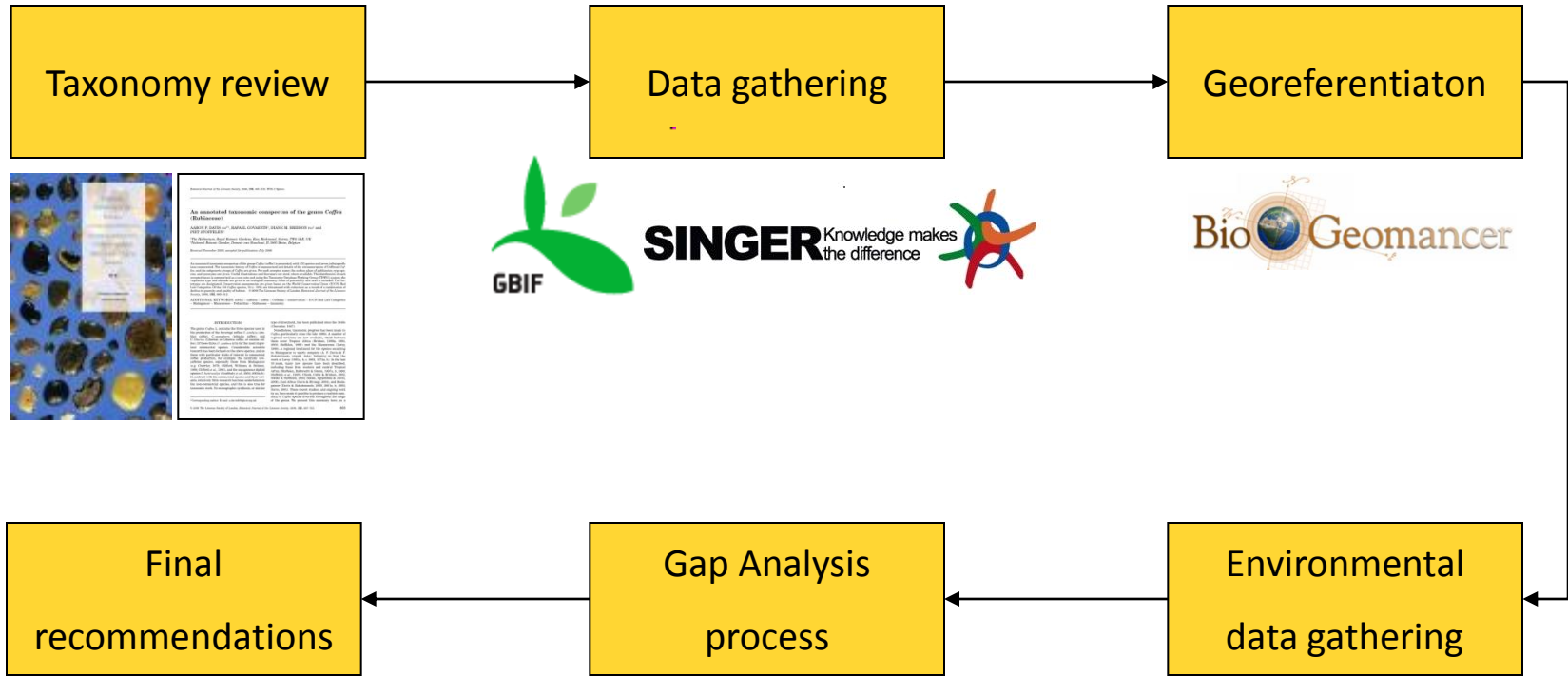
- Identifying gaps





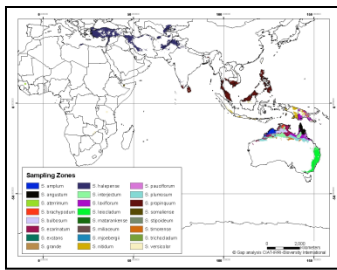


# The Gap Analysis pathway



List of priorities for the taxonomic group

Species	Number of records	Number of specimens	Number of photos	Number of drawings	Number of maps	Number of audio	Number of video	Number of other	Number of total
1. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
2. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
3. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
4. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
5. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
6. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
7. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
8. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
9. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
10. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
11. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
12. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
13. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
14. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
15. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
16. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
17. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
18. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
19. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
20. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
21. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
22. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
23. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
24. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
25. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
26. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
27. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
28. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
29. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
30. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
31. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
32. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
33. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
34. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
35. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
36. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
37. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
38. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
39. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
40. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
41. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
42. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
43. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
44. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
45. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
46. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
47. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
48. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
49. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10
50. <i>Cyrtus</i>	10	10	10	10	10	10	10	10	10

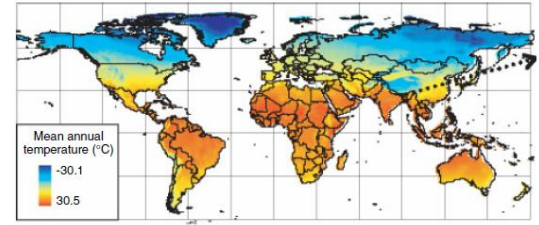


```

C:\WINDOWS\system32\cmd.exe - arc
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.
C:\Documents and Settings\NCSA\arc>
Copyright (C) 1982-2006 Environmental Systems Research Institute, Inc.
All rights reserved.
ARC 9.0.0 Sun Sep 17 16:05:34 PDT 2006

This software is provided with RESTRICTED AND LIMITED RIGHTS. Use,
Application, and disclosure by the U.S. Government are subject to
restrictions as set forth in FAR Section 52.227-14 Alternative II (c)(3)
(JUN 1987), FAR Section 52.227-19 (JUN 1987), and/or FAR Section
12.211-2 (23 Jul 1995) (Commercial Technical Data-Computer Software) and DFARS
Section 252.227-7015 (NOV 1995) (Technical Data and/or DFARS Section
252.227-7016 (Computer Software), as applicable. Contractor/Manufacturer is
Environmental Systems Research Institute, Inc., 380 New York Street,
Bedford, MA 02822-8100, USA.

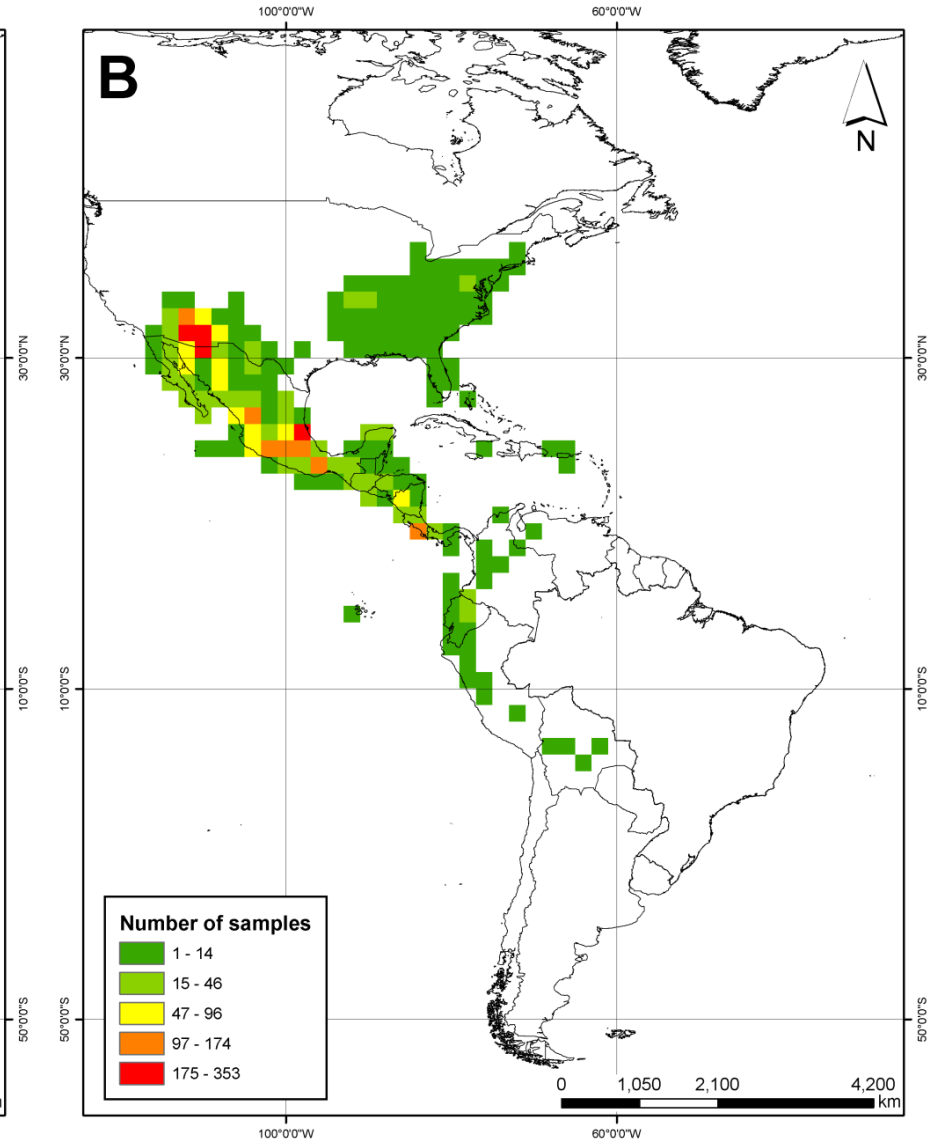
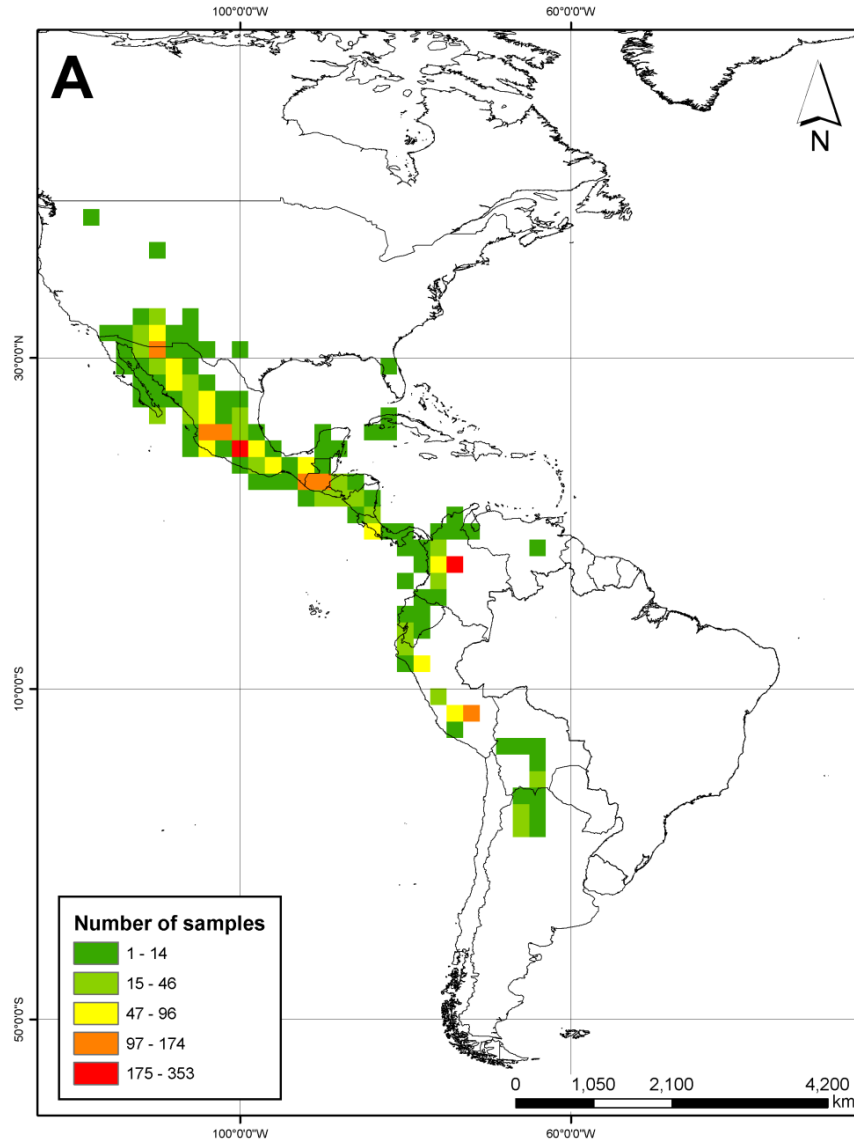
arc> _
  
```



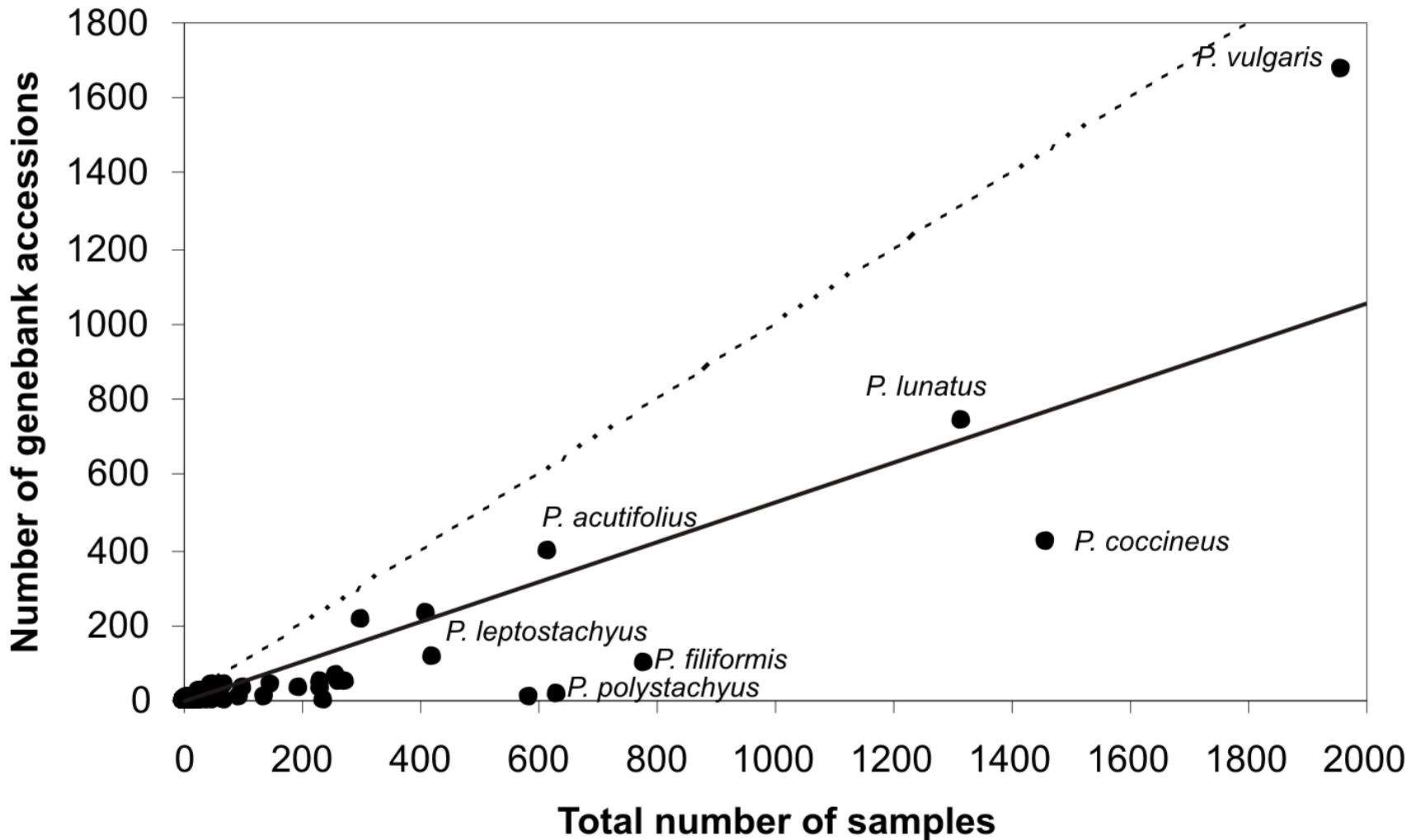
# An example in *Phaseolus*



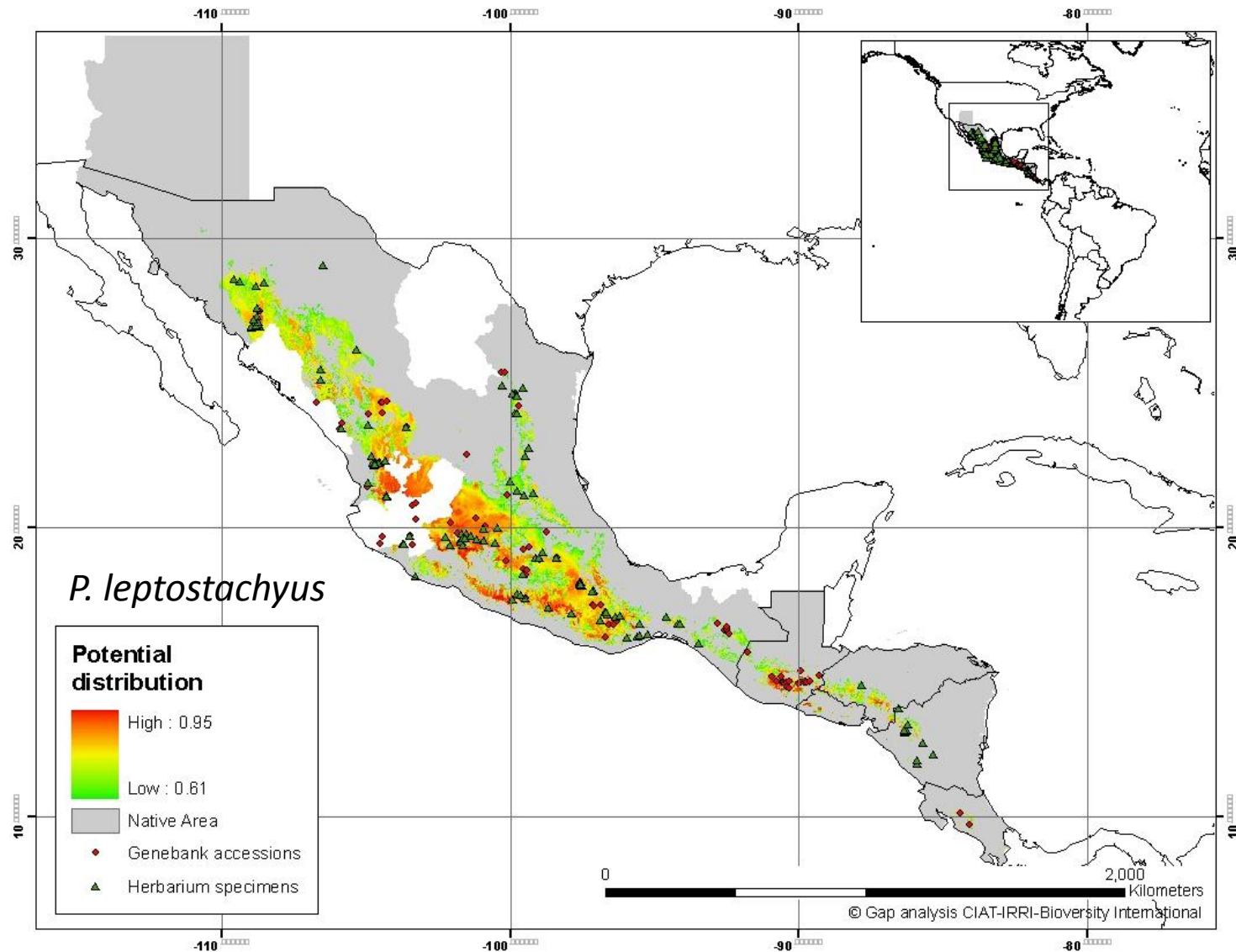
# Herbarium versus germplasm: Geographic



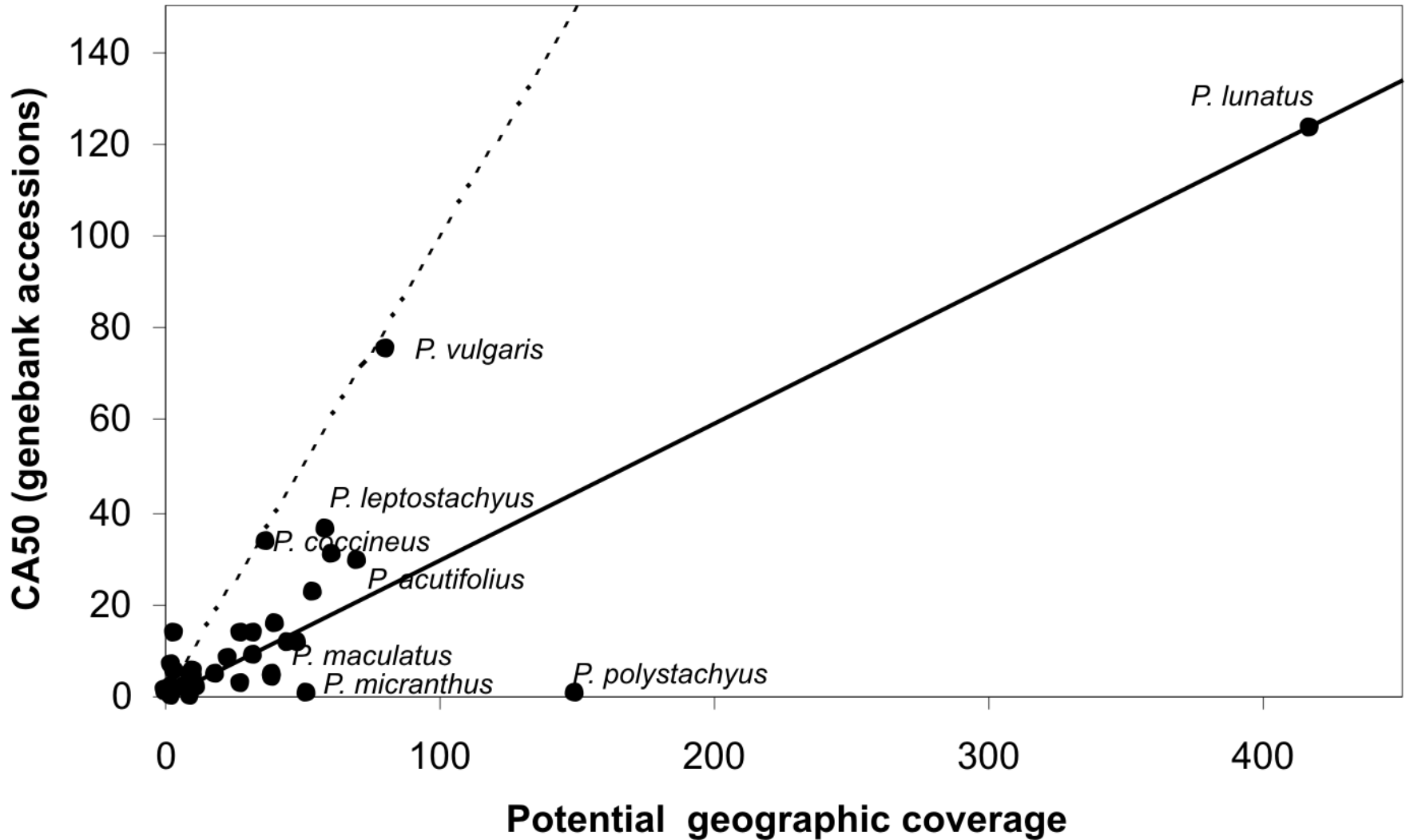
# Herbarium versus germplasm: Taxon



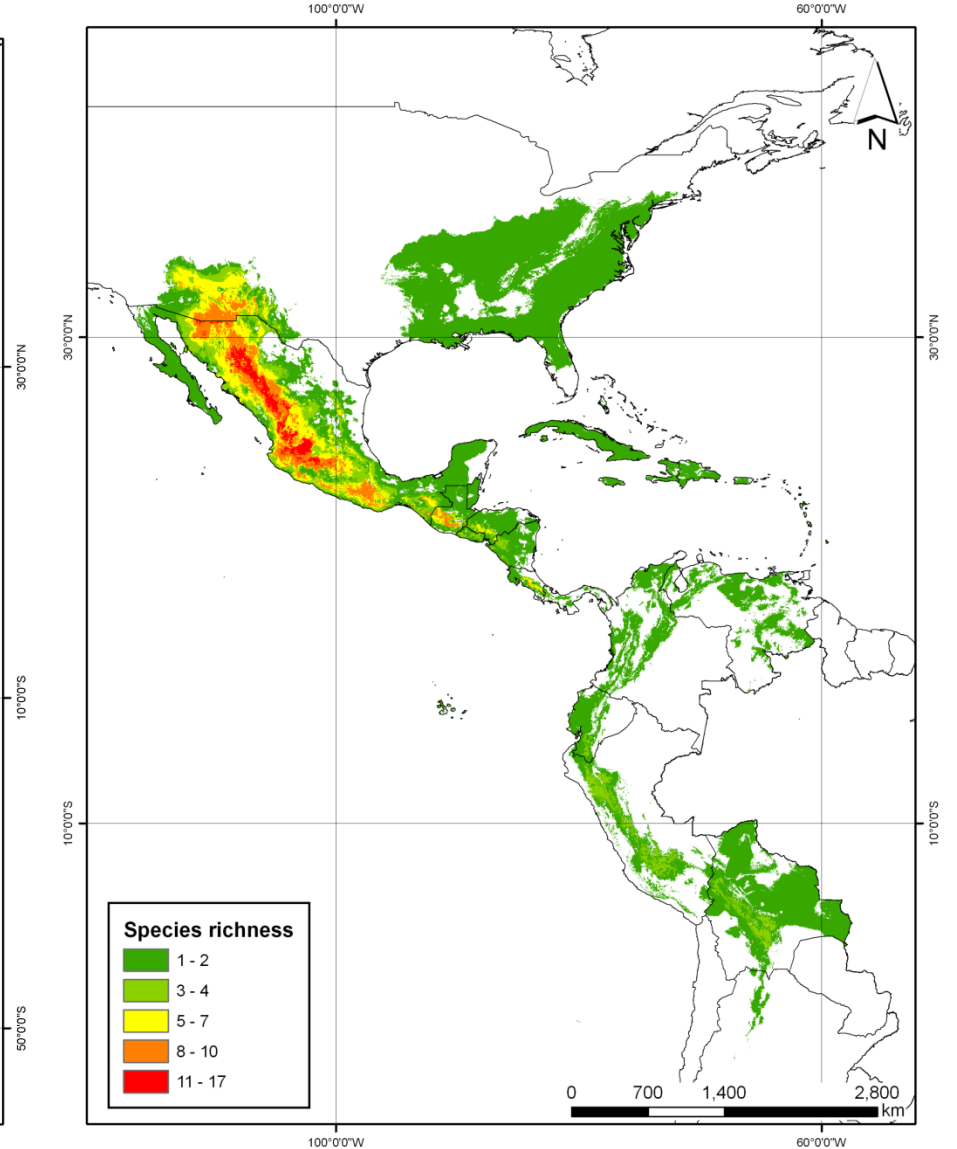
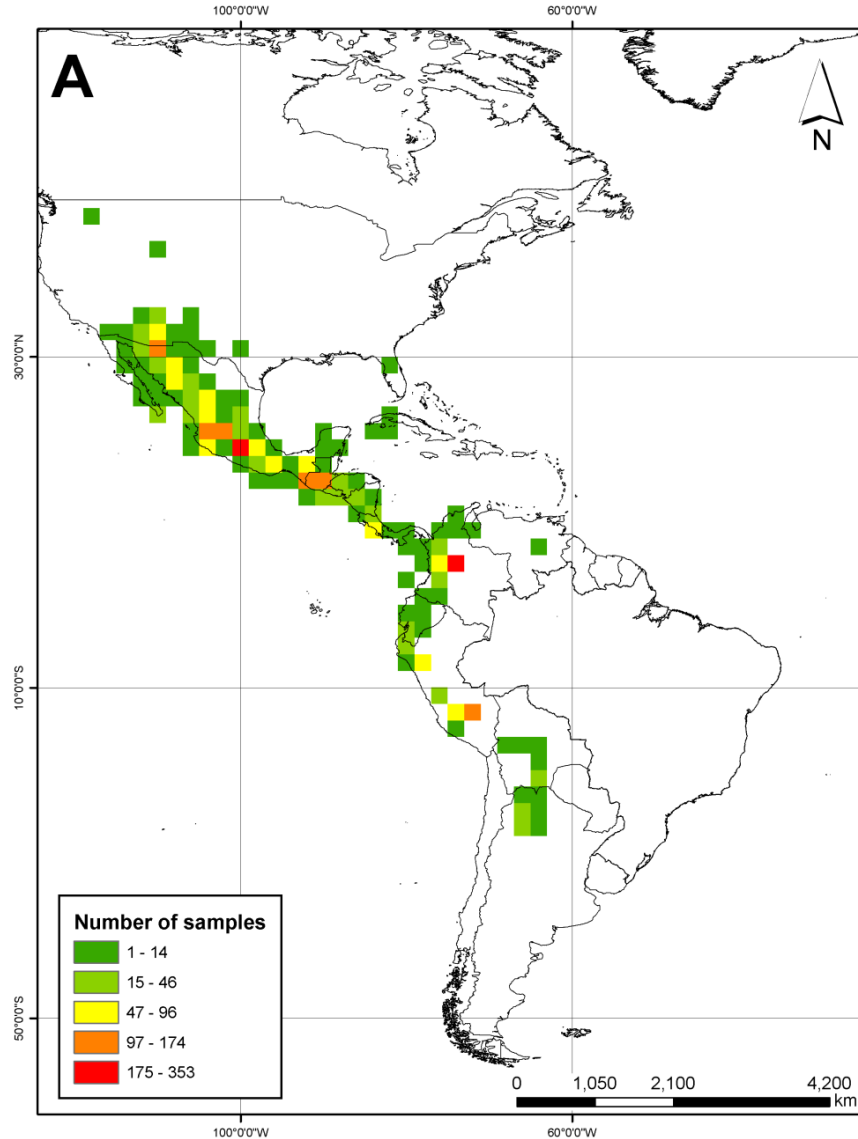
# Geographic coverage: Potential



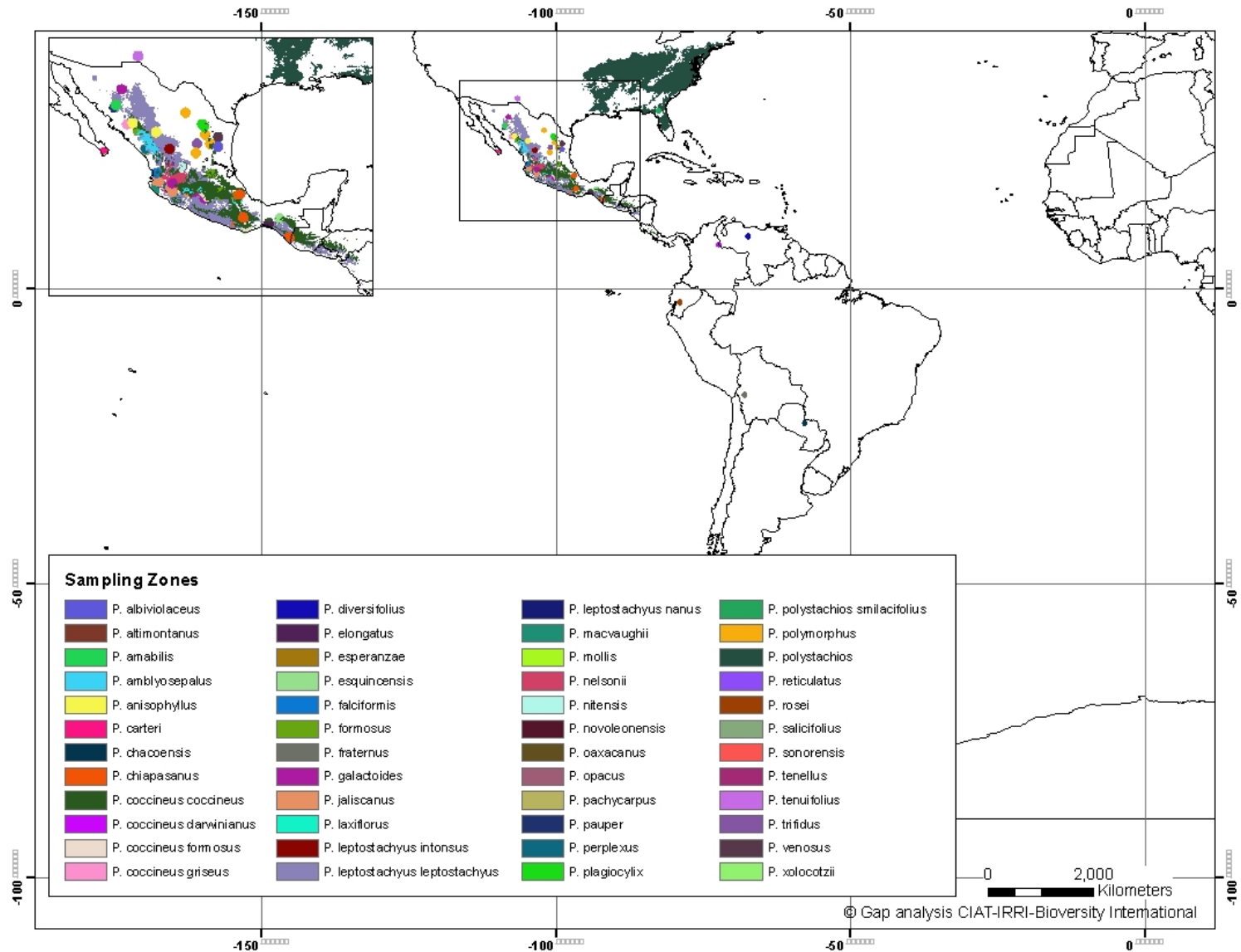
# Geographic coverage: Taxon priorities



# Conserved ex situ richness versus potential



# Priorities: Geographic and taxonomic



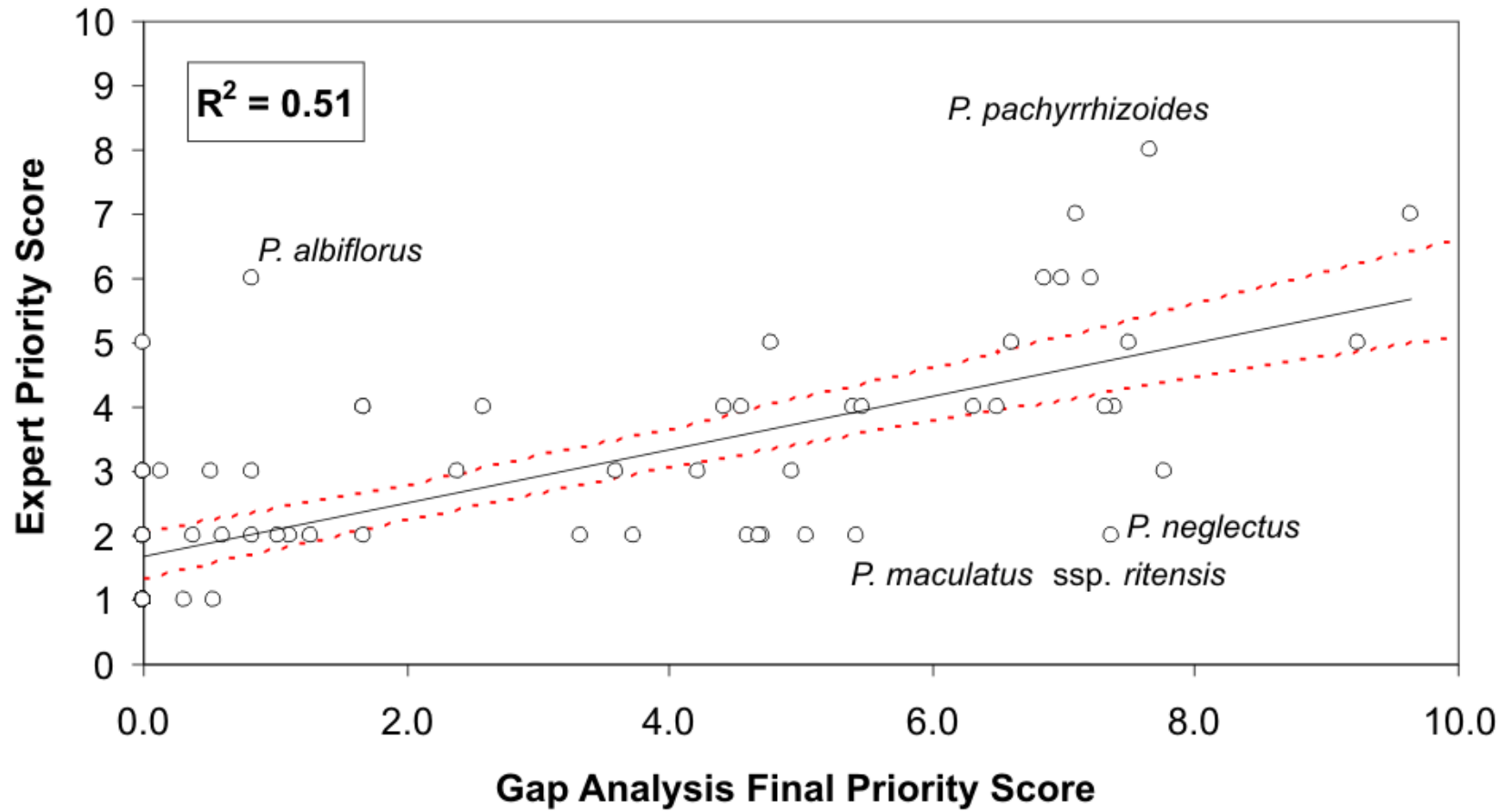


# “Validation”: The man versus the machine



# Model priorities versus expert priorities

**B**



# Gap Analysis

Home CropWildRelatives Crops Methods About us

Search

## GapAnalysis

### 2 Comments to "Phaseolus genepool"

**Jorge Berny says:**  
21 July 2009 at 12:42 PM

Hi,

I'm Jorge Berny. A rese  
Vigna. I'll be starting a G  
you know that there is gr  
and also further collectio

With regards,

Jorge C. Berny

**Julian says:**  
5 August 2009 at 9:47 AM

Great to hear that Jorge  
for us), any characterizat

As we are also intereste  
will look forward for any  
you are lack of lat/lon da  
data back.

Let us know how do you

### Leave your Comment

Submit

- [JPEG](#) Climatic niche model for *A. bicornis*
- [JPEG](#) Climatic niche model for *A. biuncialis*
- [JPEG](#) Climatic niche model for *A. caudata*
- [JPEG](#) Climatic niche model for *A. colummaris*
- [JPEG](#) Climatic niche model for *A. comosa*
- [JPEG](#) Climatic niche model for *A. crassa*
- [JPEG](#) Climatic niche model for *A. cylindrica*
- [JPEG](#) Climatic niche model for *A. kotschyi*
- [JPEG](#) Climatic niche model for *A. longissima*
- [JPEG](#) Climatic niche model for *A. neglecta*
- [JPEG](#) Climatic niche model for *A. searsii*
- [JPEG](#) Climatic niche model for *A. sharonensis*
- [JPEG](#) Climatic niche model for *A. speltooides*
- [JPEG](#) Climatic niche model for *A. tauschii*
- [JPEG](#) Climatic niche model for *A. triuncialis*
- [JPEG](#) Climatic niche model for *A. umbellulata*
- [JPEG](#) Climatic niche model for *A. uniaristata*
- [JPEG](#) Climatic niche model for *A. vavilovii*
- [JPEG](#) Climatic niche model for *A. ventricosa*
- [JPEG](#) Climatic niche model for *T. turgidum* subsp. *dicoccon*
- [JPEG](#) Climatic niche model for *T. turgidum* subsp. *dicoccoides*
- [JPEG](#) Climatic niche model for *T. urartu*
- [JPEG](#) Predicted species richness under current climatic conditions
- [JPEG](#) Predicted species richness under future climatic conditions
- [JPEG](#) Changes on predicted species richness due to climate change
- [JPEG](#) Sampling density (200km cell size) for germplasm accessions
- [JPEG](#) Sampling density (200km cell size) for herbarium samples
- [PDF](#) Conservation priorities list for *Triticum* and *Aegilops* genepool
- [XLS](#) Dataset used for this analysis



aris and Lunatus) and  
s. Anyway, just to let  
sions in our collection



ort data (which is key

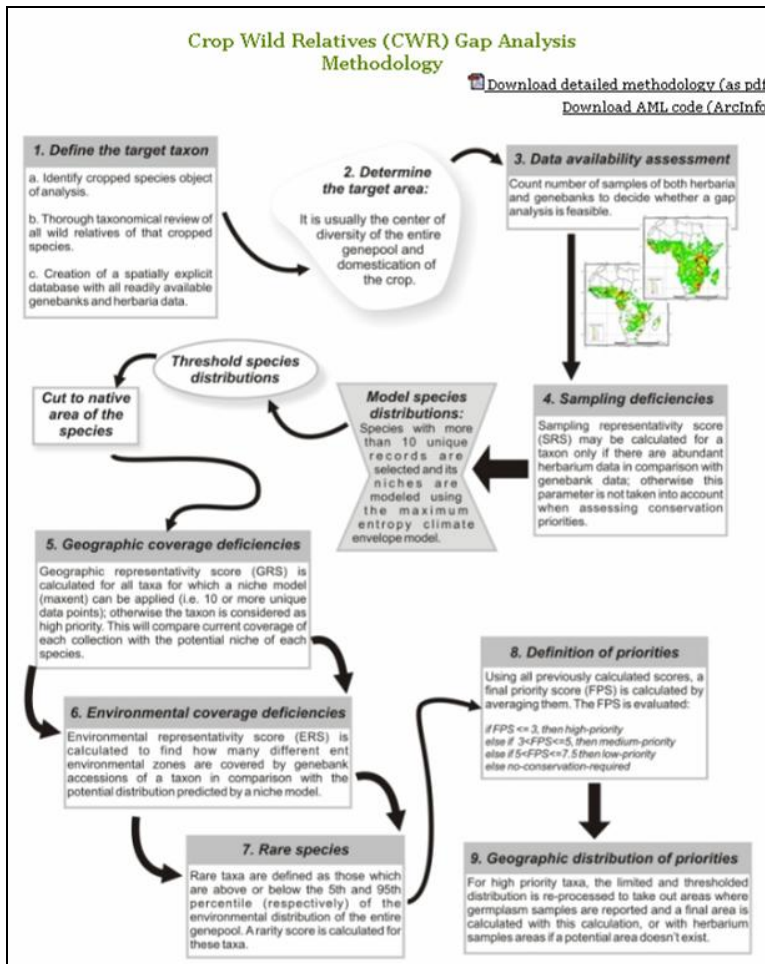
iate your interest and  
correction tools, so, if  
and then give you the

pp-right of the page).

✓ <http://gisweb.ciat.cgiar.org/GapAnalysis/>

# Methodology and code

## Methodology



## Code

```

/* This aml perform the whole gap analysis procedure for a series of genepools. You just need to change the path to the variable %rootdir%
/* and the names of the genepools to which you want to perform the analyses, and run it under the ARC prompt.
/* by Julian Ramirez-V, 2009

setlocal %%%
%*% %*%
message %off

%*%
%*% -----
%*% -- GAP ANALYSIS FOR WILD RELATIVES OF CROPPED SPECIES --
%*% - International Centre for Tropical Agriculture (CIAT) -
%*% ----- written by Julian Ramirez-V. -----
%*% ----- 2009 -----
%*%
%*%
%*% rootdir %g:\gap_analysis_v2
%*% mkdir %rootdir%\work
%*% input_dir %rootdir%\global_climate_data
%*% output_dir %rootdir%\final_output_3_2_19

%*% generalist_clocr /*%obson,clocr,alavuson,hordeum,leuca,penicillium,phaseolus,norghum,trit,maizillope,vicia,vigna,ses
%*% %spes (%generalist -count)

do i = 1 to %genus%
%*% %genusname (extract %i% %generalist%)
%*% %genuslocr (%input %genusname% 1 0)

%*%
%*% %*% Starting the process for the genepool %genusname%
%*% %genuslocr %rootdir%\gap_%genusname%
%*% %bigovoid %genuslocr%\%genusname%\_data

/* 0. SETTING UP STUFF FOR STARTING THE PROCESS
%*%
%*% %*% Preparing environmental layers
%*%
%*%
%*% %*% %rootdir%\_mask%\%genusname%
%*%
%*% %out_dir_root %genuslocr%\climate_data
%*% %out_dir_ext %out_dir_root%\current
%*% %out_dir_2050 %out_dir_root%\2050

%*% %if not (%exists %out_dir_root%\_disc) then %do %do %out_dir_root%
%*% %if not (%exists %out_dir_root%\_workspace) then %do %do %out_dir_root%
%*% %if not (%exists %out_dir_2050%\_workspace) then %do %do %out_dir_2050%

/* 0.1. MASKING ENVIRONMENTAL LAYERS AND CREATING ARCIDS
%*%
%*%
%*% GRID
%*%
%*% %*% Masking bioclimatic

do o = 1 to 19
%*%
%*% %*% %*%

```

```

C:\WINDOWS\system32\cmd.exe - arc
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

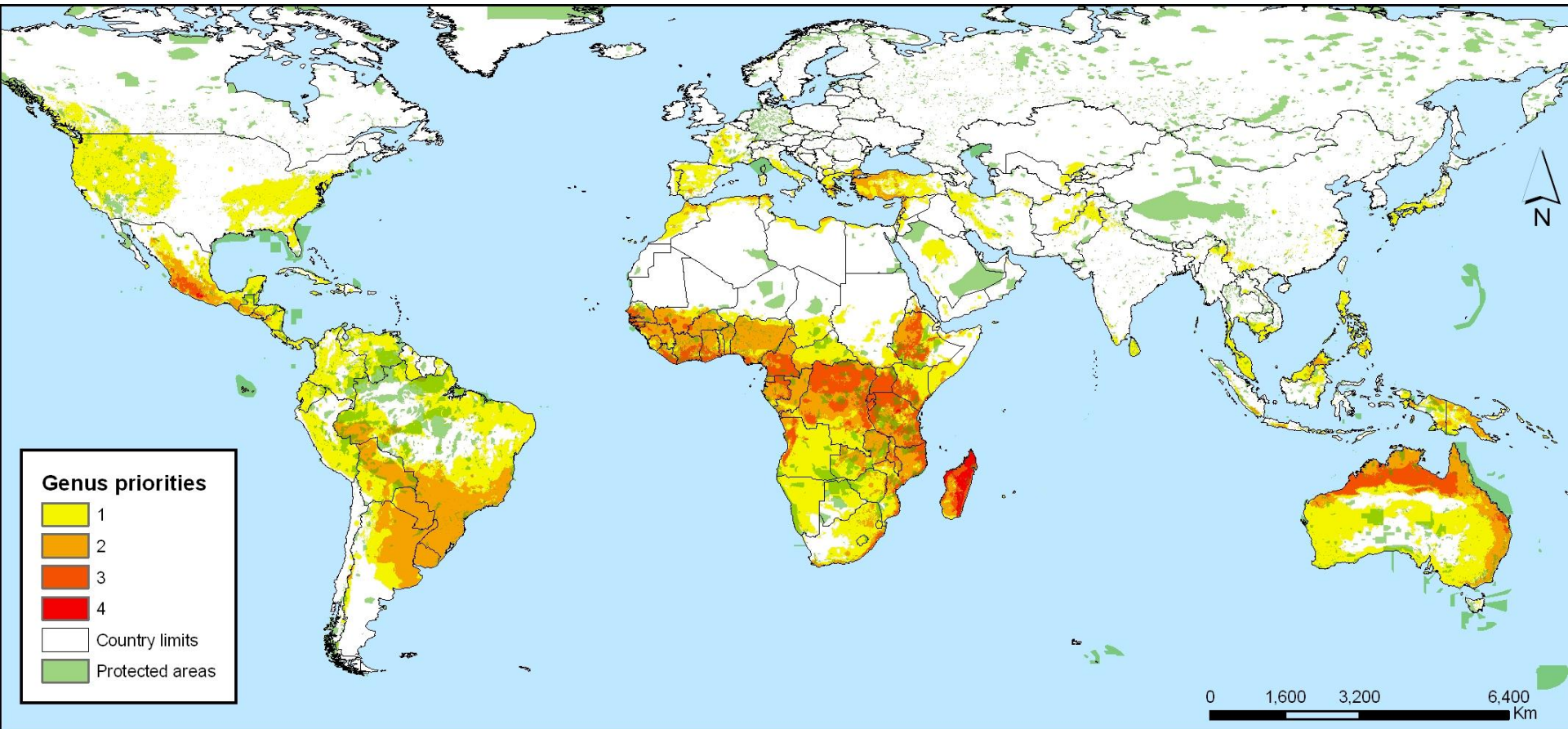
C:\Documents and Settings\ncastaneda\arc
Copyright (C) 1982-2006 Environmental Systems Research Institute, Inc.
All rights reserved.
ARC 9.2 (Sun Sep 17 16:05:34 PDT 2006)

This software is provided with RESTRICTED AND LIMITED RIGHTS. Use,
duplication, and disclosure by the U.S. Government are subject to
restrictions as set forth in FAR Section 52.227-14 Alternate III (g)(3)
(JUN 1987), FAR Section 52.227-19 (JUN 1987), and/or FAR Section
12.211/12.212 [Commercial Technical Data/Computer Software] and DFARS
Section 252.227-7015 (NOV 1995) [Technical Data and/or DFARS Section
227.7202 [Computer Software], as applicable. Contractor/Manufacturer is
Environmental Systems Research Institute, Inc., 380 New York Street,
Redlands, CA 92373-8100, USA.

Arc: -

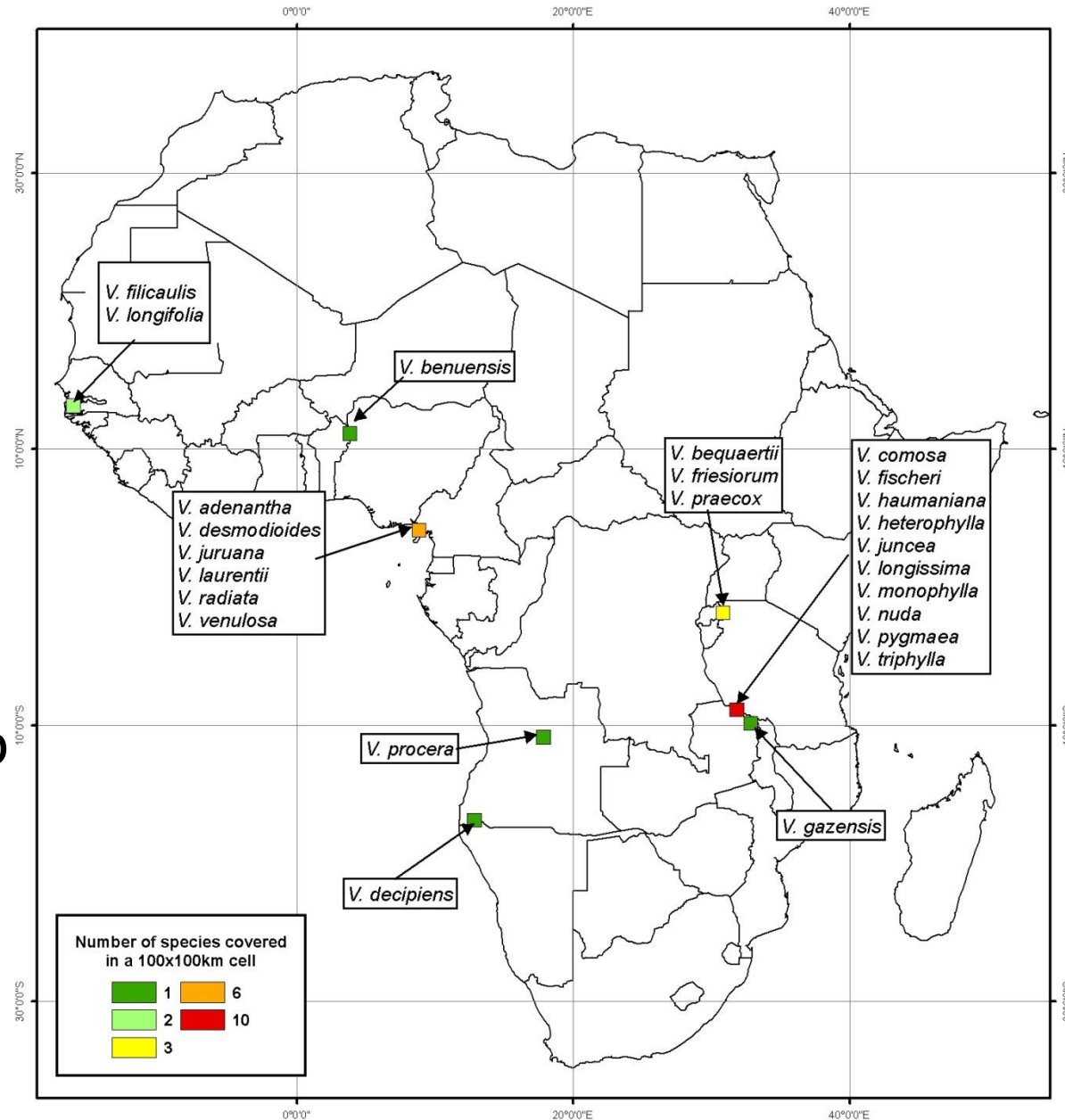
```

# Taxon-level and genepool level priorities



# Wild *Vigna* collecting priorities

- Spatial analysis on current conserved materials
- **\*Gaps\*** in current collections
- Definition and prioritisation of collecting areas
- 8 100x100km cells to complete collections of 23 wild *Vigna* priority species



**miRCURY LNA™ Universal RT microRNA PCR**  
Click here Get started now

**EXIQON**

**PLOS Currents**  
Produced with support from CHDI Foundation

New Section:  
**Huntington Disease**  
<http://currents.plos.org/hd>



Search articles...  Advanced Search

[Home](#) [Browse Articles](#) [About](#) [For Readers](#) [For Authors and Reviewers](#)

[Journals](#) [Hubs](#) [PLOS.org](#)

RESEARCH ARTICLE

OPEN ACCESS

# A Geospatial Modelling Approach Integrating Archaeobotany and Genetics to Trace the Origin and Dispersal of Domesticated Plants

| [Citation](#) | [XML](#)  
  
 New & improved!

[Article](#) [Metrics](#) [Related Content](#) [Comments: 0](#)

Jacob van Etten<sup>1,2\*</sup>, Robert J. Hijmans<sup>1,3</sup>

**1** International Rice Research Institute, Los Baños, Philippines, **2** IE University, Segovia and Madrid, Spain, **3** Department of Environmental Science and Policy, University of California Davis, Davis, California, United States of America

To add a note, highlight some text. [Hide notes](#)  
 [Make a general comment](#)

- Jump to**
- [Abstract](#)
  - [Introduction](#)
  - [Methods](#)
  - [Results](#)
  - [Discussion](#)
  - [Supporting Information](#)

**Metrics** ⓘ

**Total Article Views:** [527](#)

**Average Rating** (0 User Ratings)  
☆☆☆☆☆ [See all categories](#)  
[Rate This Article](#)

**Related Content**

**Related Subject Categories**  
[Genetics and Genomics](#), [Ecology](#), [Plant Biology](#)

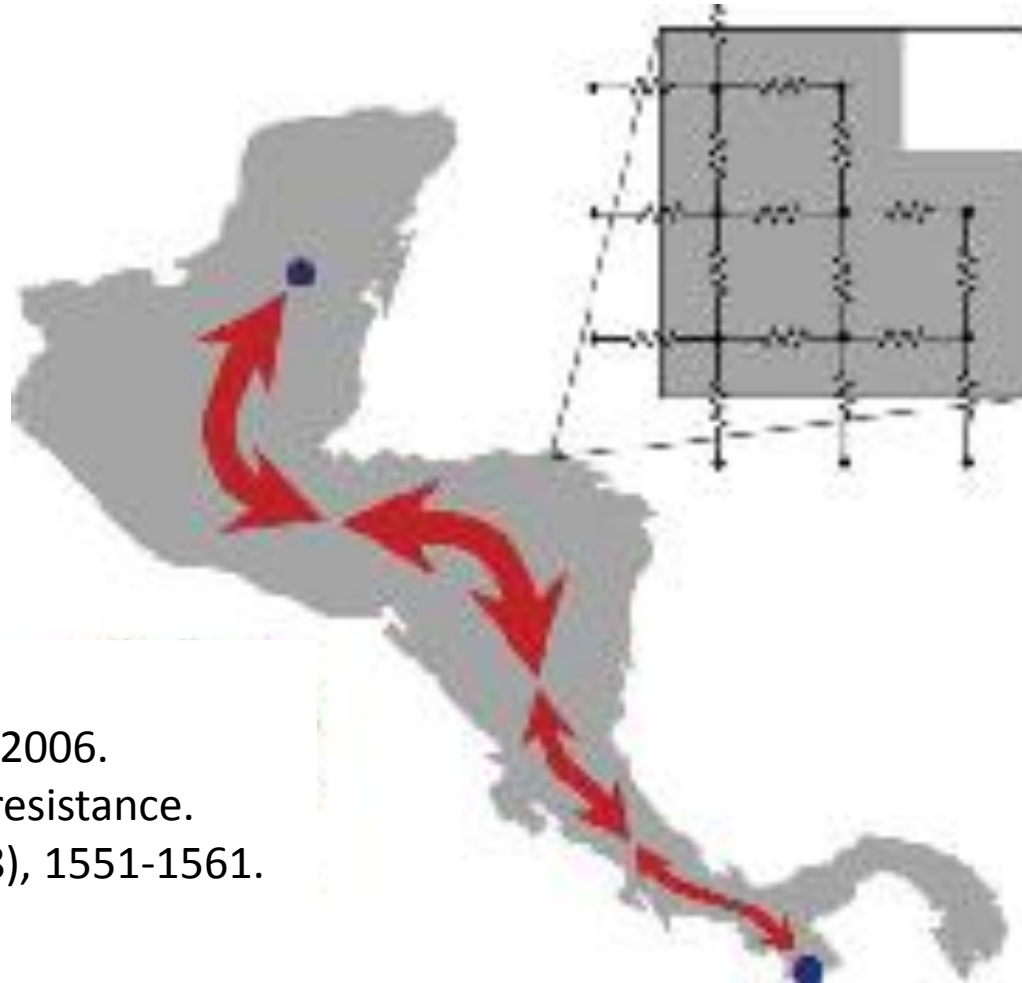
**Related Articles on the Web**  
[Google Scholar](#)

**Abstract** [Top](#)

**Background**

The study of the prehistoric origins and dispersal routes of

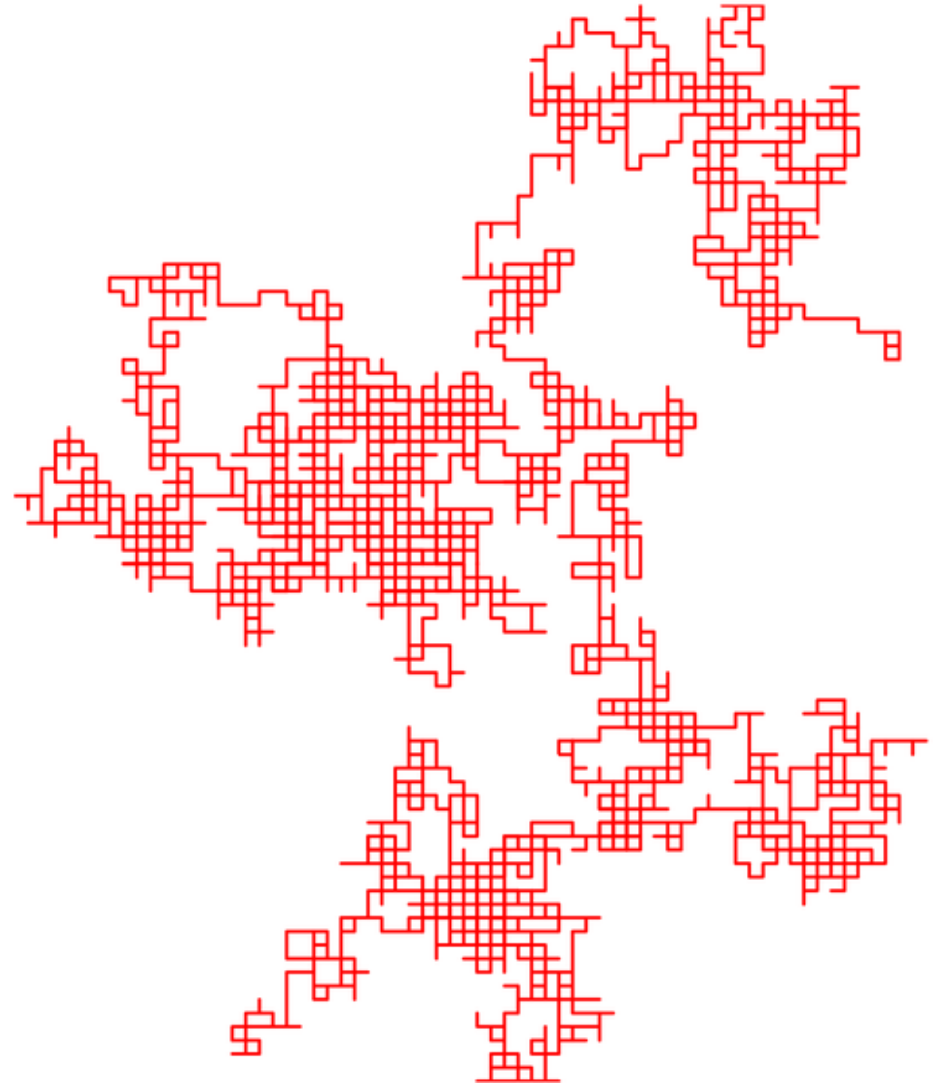
# Even better: resistance distance



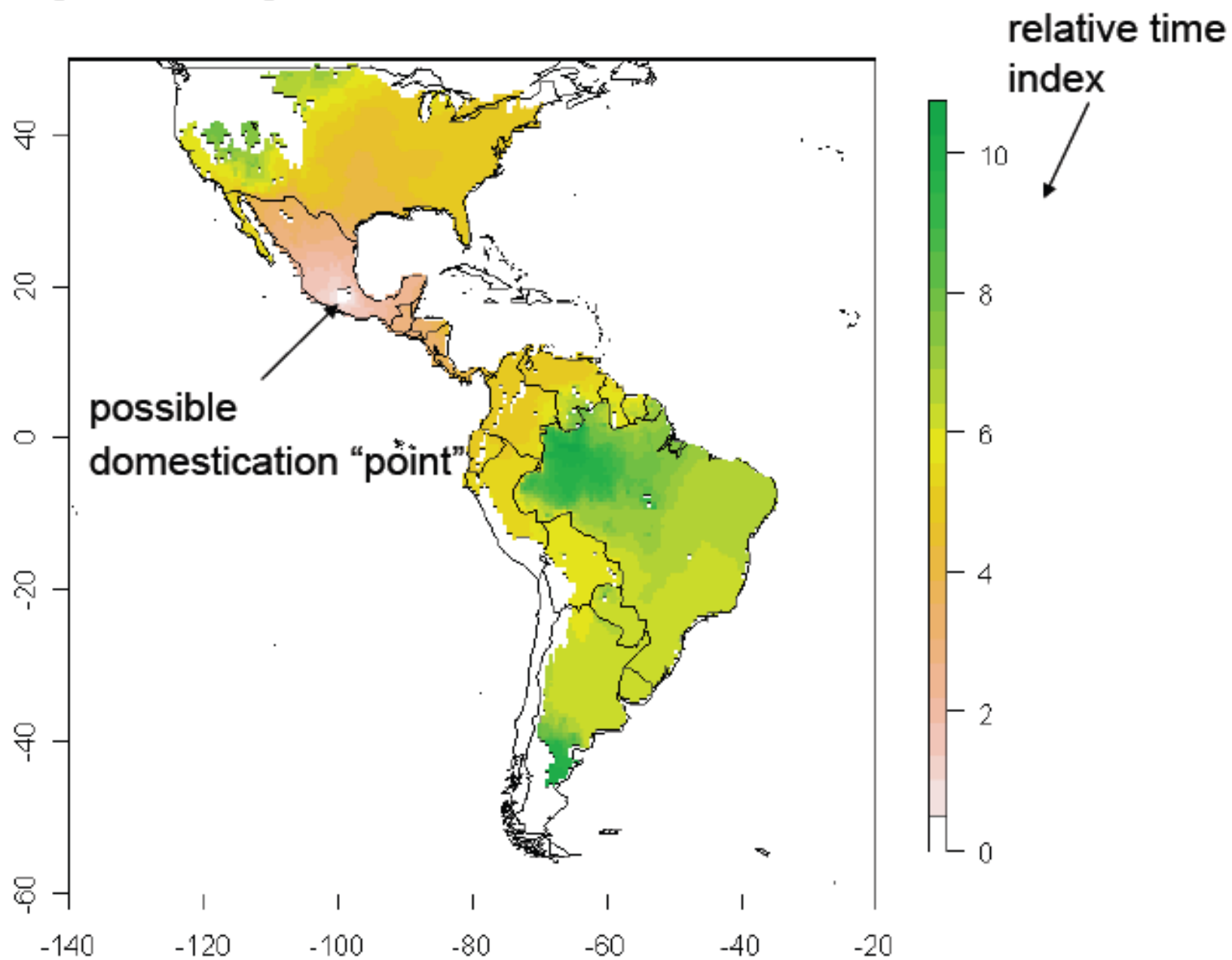
McRae, B.H. 2006.  
Isolation by resistance.  
*Evolution* 6(8), 1551-1561.



# Drunk genes



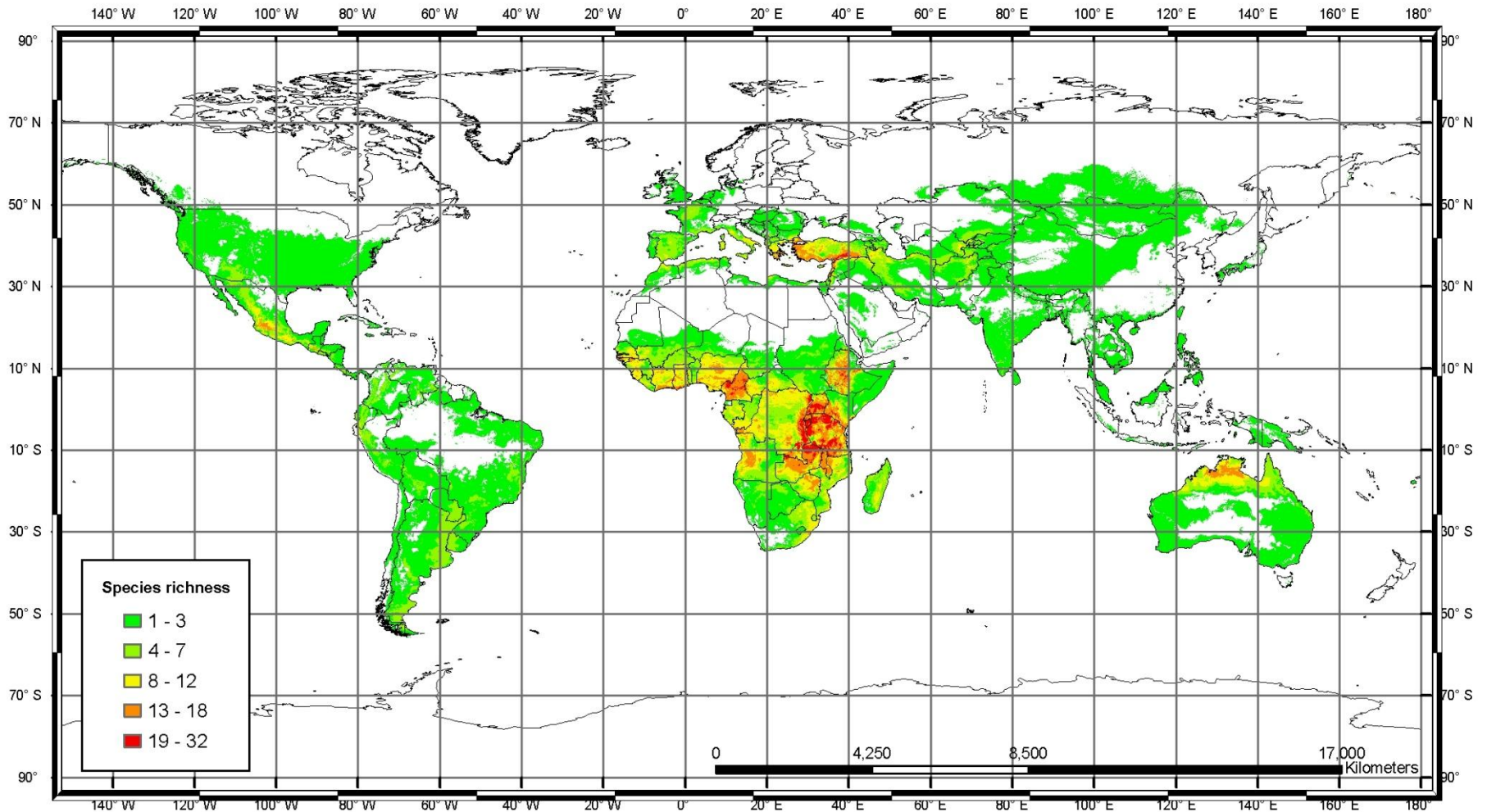
# Crop dispersal arrival time



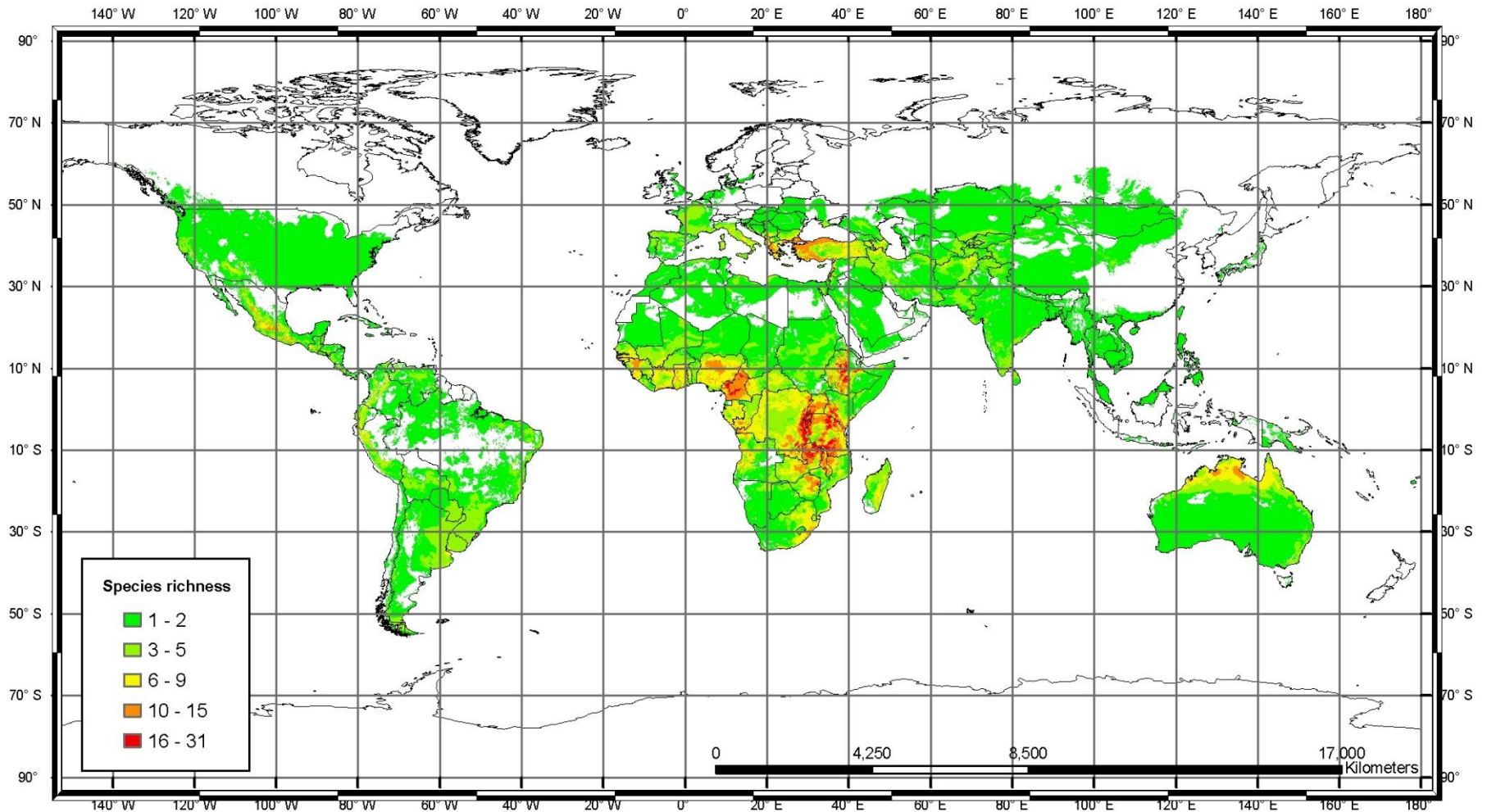


Threats

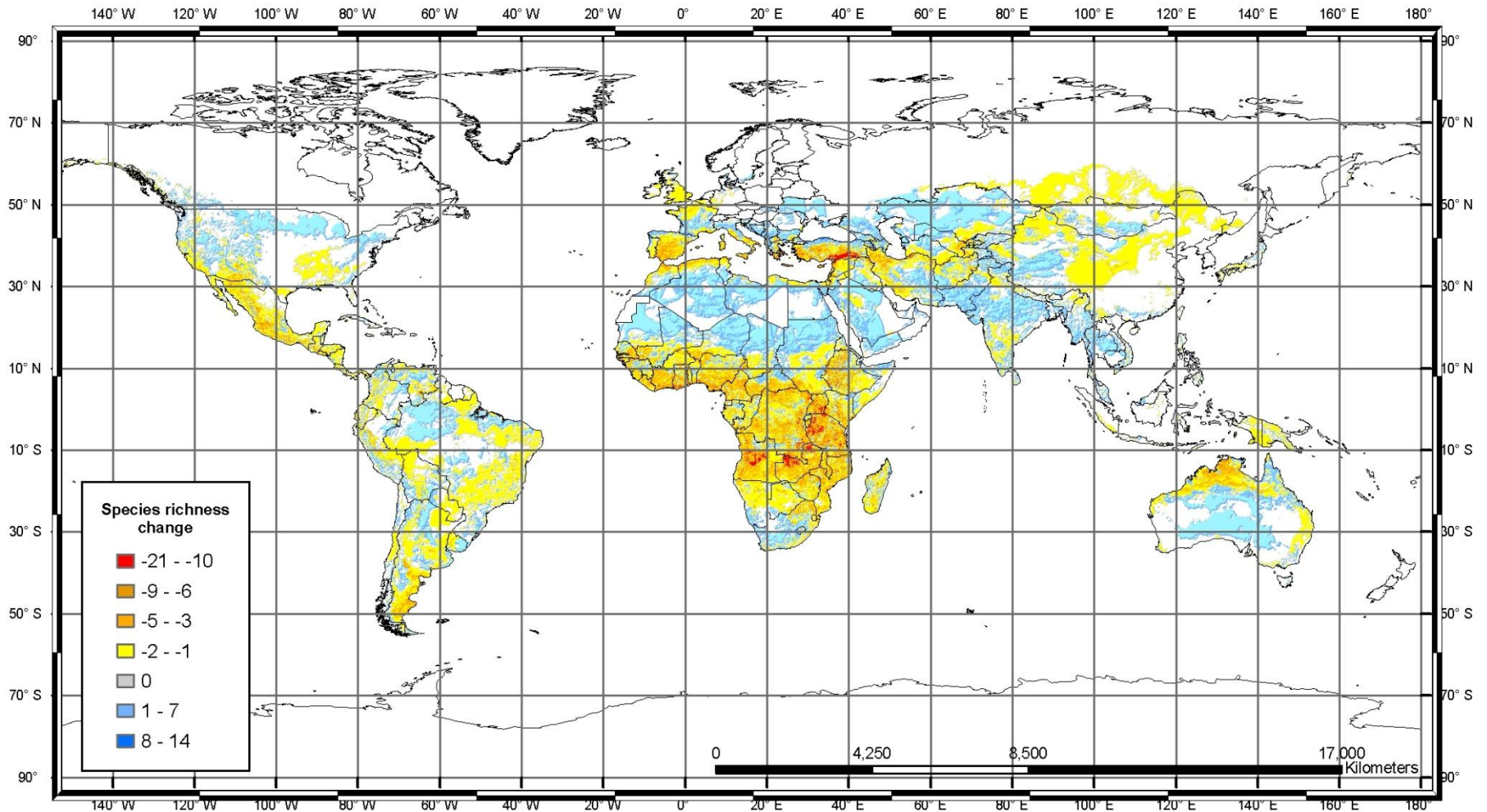
# CWR supporting adaptation but also threatened by climate change



# CWR supporting adaptation but also threatened by climate change



# CWR supporting adaptation but also threatened by climate change



More immediate threats....

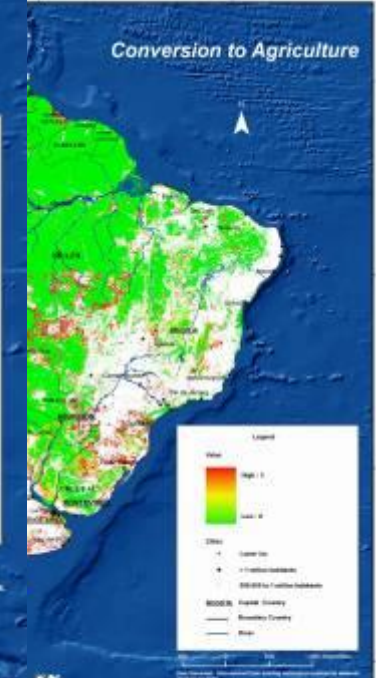
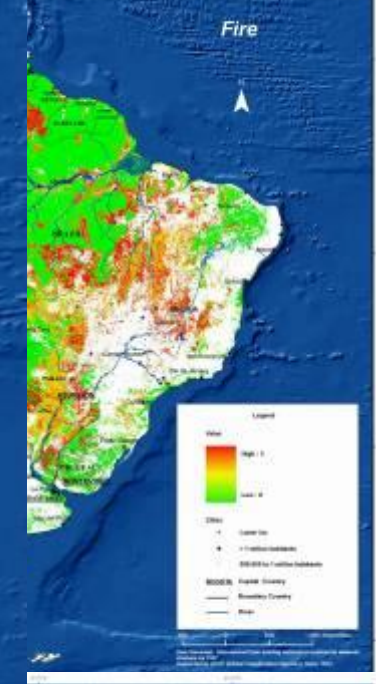
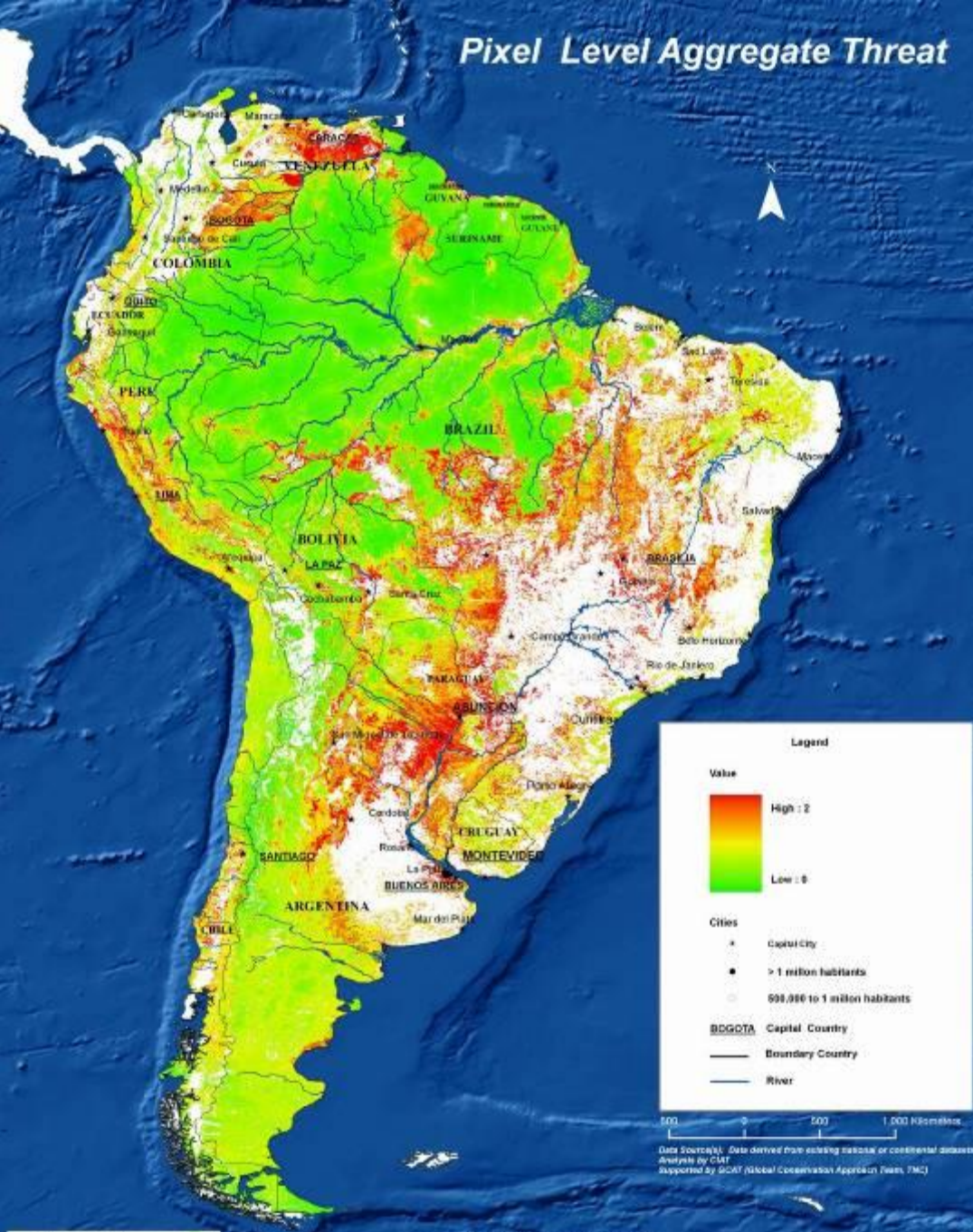
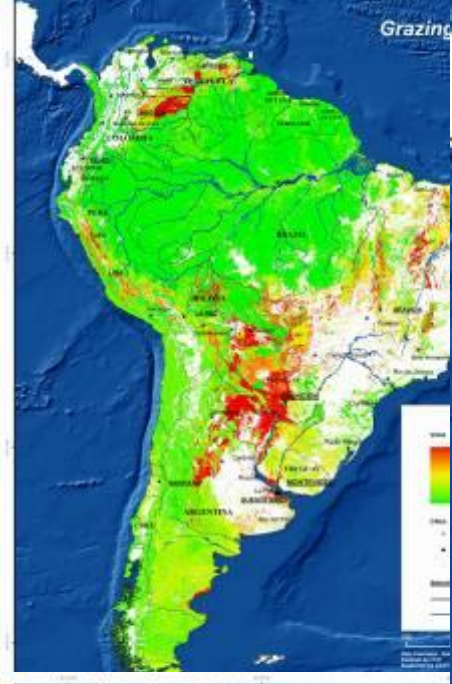
Grazing

# Pixel Level Aggregate Threat

Fire

Oil a

Conversion to Agriculture



**Legend**

Value

High : 2

Low : 0

Cities

- Capital City
- > 1 million habitants
- 500,000 to 1 million habitants

BDGOTA Capital Country

Boundary Country

River

500 1,000 Kilometers

Data Source(s): Data derived from existing national or continental datasets.  
 Analysis by CIAT  
 Supported by CIAT (Global Conservation Approach Team, FMC)



# Conclusions and open-ended issues

- Gap analysis can focus priorities for collecting, and cost/benefit lies in collecting multiple genepools simultaneously
- But, analysis as good as input data
- If you can't see the data, is the collection valid?
- Multiple methods available
- Key question: what is a “complete” collection?