Bretting, Peter

From:	Peter Wenzl <peter.wenzl@divseek.org>
Sent:	Thursday, September 17, 2015 4:37 AM
To:	Andreas Graner; David Marshall; Emily Marden; Bretting, Peter; Rajeev Varshney (ICRISAT-IN); Ruaraidh Sackville Hamilton (IRRI); Sarah Ayling; Susan McCouch
Cc:	Daniele Manzella; Ruth Bastow; Wayne Powell (CGIAR Consortium); Pascale (CGIAR Consortium)
Subject:	DivSeek Steering Committee meeting in Bonn on Dec 8th

Dear Steering Committee members,

After consultations with Susan and among JFU members, we'd like to propose to hold the next DivSeek Steering Committee meeting in Bonn on December 8th at the premises of the Crop Trust in Bonn.

Suggested agenda items for the meeting :

- Review first report from the governance committee
- Review the widening landscape study
- Review new membership applications
- Prepare for the upcoming DivSeek Partners' Assembly (Jan. 2016):
  - Draft a Program of Work
  - Draft a Resource Mobilization Plan
- Discuss a possible Bellagio conference focused on the interface between PGR-related science opportunities and associated policy challenges

Please let us know whether will be able to come to Bonn (in person) for this meeting and whether there are any additional agenda items you would like to discuss.

Best regards,

Peter (on behalf of the JFU)

--
Peter Wenzl
DivSeek Liaison
Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted] (b) (6)
www.croptrust.org

Securing our Food, Forever

Bretting, Peter

From: Peter Wenzl <peter.wenzl@croptrust.org>
Sent: Saturday, August 29, 2015 5:43 AM
To: Bretting, Peter
Cc: Daniele Manzella (ITPGRFA); Ruth Bastow (GPC); Susan McCouch; Powell, Wayne (CGIAR Consortium)
Subject: Re: FW: Outline for a DivSeek Knowledge Exchange and Capacity-Building Workshop

Dear Peter,

Many thanks for sharing this proposal! I believe it deals with a number of topics that are central to DivSeek. While reading it, I kept asking myself whether starting with the "big picture" or with a limited task (such as a common standard for GbS data) would be more effective. I guess there're pros and cons in both directions, and perhaps we need both, albeit for different reasons.

Best, Peter

On Tue, Aug 25, 2015 at 7:36 PM, Bretting, Peter <Peter.Bretting@ars.usda.gov> wrote:

Hi Peter, Daniele, and Ruth—apologies, I forgot to address a copy to you when writing to Susan.

Thanks, hope that you have enjoyed a pleasant and productive summer!

Peter

Peter Bretting

USDA/ARS Office of National Programs

Room 4-2212, Mailstop 5139

5601 Sunnyside Avenue

Beltsville, MD 20705-5139

Phone 1.301.504.5541

Fax 1.301.504.6191

Mobile Phone [redacted]

E-mail peter.bretting@ars.usda.gov

Web site: http://www.ars.usda.gov/research/programs/programs.htm?NP_CODE=301

From: Bretting, Peter
Sent: Tuesday, August 18, 2015 6:54 AM
To: 'Susan McCouch'; 'Susan McCouch'
Subject: Outline for a DivSeek Knowledge Exchange and Capacity-Building Workshop

Hi Susan—paragraph 29 of the DivSeek Steering Committee meeting report states: "To further develop point d) above, the Chairperson requested that each Committee member draft a one-page outline for a DivSeek knowledge exchange and capacity building workshop."

Attached is my homework! It's the result of discussions with Carson Andorf (Maize GDB), Steven Cannon (Soybase and Legume Information System), and Chris Richards (NCGRP). They are responsible for any incisive thinking, whereas I'm responsible for any errors.

Thanks,

Peter

Peter Bretting

USDA/ARS Office of National Programs

Room 4-2212, Mailstop 5139

5601 Sunnyside Avenue

Beltsville, MD 20705-5139

Phone 1.301.504.5541

Fax 1.301.504.6191

Mobile Phone [redacted]

E-mail peter.bretting@ars.usda.gov

Web site: http://www.ars.usda.gov/research/programs/programs.htm?NP_CODE=301

--

Peter Wenzl
DivSeek Liaison
Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]
www.croptrust.org

Securing our Food, Forever

Bretting, Peter

From:	Peter Wenzl <peter.wenzl@croptrust.org>
Sent:	Tuesday, May 12, 2015 3:14 PM
To:	Nelson, Randall - ARS
Cc:	Scott Allen Jackson; Cannon, Steven; Suk-Ha Lee; Jones Jennifer; Abberton Michael (IITA); Cannon Steven [AGRON]; Joost Richard; Okamuro, Jack; 邱丽娟 QIU, Lijuan; 李英慧; Masao Ishimoto; Akito Kaga; Bob Stupar; Ricardo Abdelnoor; Bretting, Peter; marcelofernandes.oliveira@embr; François Belzile; Matija Obreza; Luigi Guarino
Subject:	Re: Follow Up: Glycine meeting Seattle, April 15-17

Dear Randy et al.,

In a way Genesys was designed based on a similar principles as you as a group have decided to proceed; in other words: start small with what's feasible and gradually expand into more challenging areas.

For now, Genesys can hold genebank IDs, accession IDs, passport data, and basic characterisation data. Genesys will almost certainly be expanded in the near future to link to data repositories holding more complex phenotypic data sets (e.g., field-based evaluation trials).

And in the context of DivSeek, it's only a question of time until Genesys can cross-talk to data repositories like Germinate (http://ics.hutton.ac.uk/germinate), which can already hold genotyping, phenotypic and environmental data.

I'm copying our Genesys experts (Matija, Luigi) here, so that they can clarify any questions you may have.

Kind regards, Peter

On Tue, May 12, 2015 at 7:58 PM, Nelson, Randall - ARS <Randall.Nelson@ars.usda.gov> wrote:

Scott and Peter,

Right now I am swamped with getting ready to plant and working on grant proposals so I have been putting off doing anything about these data until June.

I have heard of this database but I don't know anything about it. Peter's email states that the database is designed for "assembling global inventories of genebank accessions (accession IDs, genebank IDs & passport data)". Does it have the capacity for phenotypic and genotypic data also? If it does, I would think we would want to set up some kind of structure to make sure that the data is useable. Peter, I would be willing to do some preliminary work to explore these options if you can provide me with a contact at Genesys, but I probably won't get much done until June.

Randy

From: Scott Allen Jackson [mailto:sjackson@uga.edu]
Sent: Tuesday, May 12, 2015 12:34 PM
To: Peter Wenzl
Cc: Cannon, Steven; Suk-Ha Lee; Jones Jennifer; Abberton Michael (IITA); Cannon Steven [AGRON]; Joost Richard; Nelson, Randall - ARS; Okamuro, Jack; 邱丽娟 QIU, Lijuan; 李英慧; Masao Ishimoto; Akito Kaga; Bob Stupar; Ricardo Abdelnoor; Bretting, Peter; marcelofernandes.oliveira@embr; François Belzile

Subject: Re: Follow Up: Glycine meeting Seattle, April 15-17

This looks like a good approach for us.

Randy, are you familiar with this? Instead of you being the conduit for all the accession names/numbers, would this work?

scott

> On May 12, 2015, at 11:08 AM, Peter Wenzl <peter.wenzl@croptrust.org> wrote:
>
> Dear all,
>
> I've talked to my colleagues here at the Crop Trust who're developing and maintaining the Genesys portal (https://www.genesys-pgr.org/welcome). I hear that it would be quite straightforward to provide room for assembling a global inventory of soybean accessions on this platform, if you wish to do so.
>
> Genesys was designed for the purpose of assembling global inventories of genebank accessions (accession IDs, genebank IDs & passport data). That fact that soybean is not an Annex 1 crop of the International Treaty doesn't matter.

Best regards, Peter

On Wed, May 6, 2015 at 8:40 PM, Scott Allen Jackson <sjackson@uga.edu> wrote:

Thanks for following up Steven,

several issues here:

1) on the funding. I think you are right that initially, at least, individual but coordinated efforts will be necessary. As a group, I think we can provide letters of support stating that the proposed research is part of a larger international initiative.

2) On the collection of accessions (both genotype and non-genotyped). Randy Nelson had 'volunteered' to start the process of collecting all the accessions to begin global inventory. Perhaps, as part of that, people could indicate which ones are genotyped and how. I suspect that much of of the collections have been genotyped in the past, but probably not sequence-based (e.g. SSRs), so the informaiton may not be of much use for what we envisage.

As an aside, we had discussed having an intern work on the data analysis. But the more I think about it, I think this is going to be more of a postdoctoral project—as you indicate below.

scott

On May 1, 2015, at 12:09 PM, Cannon, Steven <Steven.Cannon@ARS.USDA.GOV> wrote:

Speaking for myself: I am still very enthusiastic about the group's plans, but have been swamped, .

I do, however, have a little progress to report. A student group that I have been advising has worked out a very fast GBS variant-calling and imputation pipeline for soybean. It is processing a lane of GBS data (384 samples multiplexed) in about two hours. This uses something like the "flanking sequencing database" concept that we discussed, to help speed the analysis and focus on known, well-assayable SNP locations. This process goes from raw reads through deconvolution of accessions through variant calling and VCF

creation and haplotype imputation. I expect they would be willing to share the methods, but after some more documentation and testing.

I have also spoken with Michelle Graham and Bill Beavis here at ISU. Michelle has resequenced about a hundred lines used in the U.S. collection, and Bill Beavis is working with Brian Diers on a genotyping of a set of NAM populations in the U.S.

A next step for my continued involvement would be to find more skilled hands (probably a postdoc). I'm still considering a couple of RFPs: one at my university, and one with the United Soybean Board. It seems to me that others in this group might also seek funds. We could try to get funding for the group – but considering the many institutions and countries involved, it may be easier for each group to try to find a small amount of funding – or otherwise carve out some time from existing staffing – and then pool our efforts.

Another relatively small but important thing that we could do with moderate effort is for each participant in the project to work to assemble a list of their country's accessions, and also a list of accessions that have been (or are being) genotyped. These lists could initially be shared via email to this group – perhaps best, to a designated person who can cross-check and integrate and summarize the lists. I hate to heap more work on Randy, but maybe his group would be in the best position to make sense of the IDs/accessions from the several countries.

Best,

Steven

Steven Cannon, PhD

USDA - Agricultural Research Service

Iowa State University

515-294-6971

From: Scott Allen Jackson <sjackson@uga.edu>
Date: Thursday, April 30, 2015 at 8:00 PM
To: Peter Wenzl <peter.wenzl@croptrust.org>

Cc: Suk-Ha Lee <sukhalee@snu.ac.kr>, Jones Jennifer <jjones@smithbucklin.com>, "Abberton Michael (IITA)" <m.abberton@cgiar.org>, Steven Cannon <scannon@iastate.edu>, Joost Richard <RJoost@smithbucklin.com>, "Nelson, Randall - ARS" <Randall.Nelson@ARS.USDA.GOV>, "Okamuro, Jack" <Jack.Okamuro@ARS.USDA.GOV>, "邱丽娟 QIU, Lijuan" <qiulijuan@caas.cn>, 李英慧 <liyinghui@caas.cn>, Masao Ishimoto <ishimoto@affrc.go.jp>, Akito Kaga <kaga@nias.affrc.go.jp>, Robert Stupar <stup0004@umn.edu>, Ricardo Abdelnoor <ricardo.abdelnoor@embrapa.br>, "Bretting, Peter" <Peter.Bretting@ARS.USDA.GOV>, "marcelofernandes.oliveira@embr" <marcelofernandes.oliveira@embrapa.br>, Steven Cannon <steven.cannon@ars.usda.gov>, François Belzile <Francois.Belzile@fsaa.ulaval.ca>
Subject: Re: Follow Up: Glycine meeting Seattle, April 15-17

This is great, Peter. I don't have any thing to add. Maybe others do. Please let us know if they have any feedback.

Thanks for taking the time to join us.

scott

> On Apr 30, 2015, at 7:15 PM, Peter Wenzl <peter.wenzl@croptrust.org> wrote:
>
> Dear all,
>
> Time's running and it's already two weeks since the meeting in Seattle! I wanted to thank you -and Scott in particular- for having invited me to your meeting. I really enjoyed it and there was a lot I leaned from your presentations and the ensuing discussions.
>
> If you don't mind, I would like to report back to the DivSeek Steering Committee some of the key take-away points from the meeting, such as:
>
> - Your plan to assemble a cross-genebank accession inventory
> - Your approach to attempt to integrate existing data sets through a community-wide catalogue of GBS sequence tags identified across a broad range of accessions
> - The plan to genotype all existing core collections on a single platform to create a global picture of the total diversity available,

which could then be used to place individual accessions into a global context
- Your plans to explore opportunities of working with DivSeek in the area of software tools, databases, and hosting large datasets.
- Your quest to at least achieve data-sharing for cases where germplasm sharing is difficult.

In case there're any other important items I've missed here, which you would like me to report back to the Steering Committee, please let me know.

Looking forward to a continuing relationship,

All the best, Peter

--

Peter Wenzl

DivSeek Liaison

Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]

www.croptrust.org

Securing our Food, Forever

On Mon, Apr 20, 2015 at 5:45 AM, "이석하" <sukhalee@snu.ac.kr> wrote:

> Dear Scott,
>
> Thanks for senind summary of our activity at Seattle. I enjoyed the meetign.

Sincerelyn yours,

Suk-Ha Lee

Department of Plant Science

College of Agriculture and Life Sciences

Seoul National University

Seoul 151-921, Korea

(Tel)82-2-880-4545 (Fax) 82-2-873-2056

---------- 원본 메일 ----------

보낸사람 : Scott Allen Jackson <sjackson@uga.edu>

받는사람 : "Jones, Jennifer" <jjones@smithbucklin.com>, "Abberton, Michael (IITA)" <m.abberton@cgiar.org>, "Cannon, Steven [AGRON]" <scannon@iastate.edu>, "Joost, Richard" <RJoost@smithbucklin.com>, "Nelson, Randall" <Randall.Nelson@ars.usda.gov>, "Okamuro, Jack" <Jack:Okamuro@ars.usda.gov>, Peter Wenzl <peter.wenzl@croptrust.org>, Suk-Ha Lee <sukhalee@snu.ac.kr>, "邱丽娟 QIU, Lijuan" <qiulijuan@caas.cn>, "李英慧" <liyinghui@caas.cn>, Masao Ishimoto <ishimoto@affrc.go.jp>, Akito Kaga <kaga@nias.affrc.go.jp>, Bob Stupar <stup0004@umn.edu>, Ricardo Abdelnoor <ricardo.abdelnoor@embrapa.br>, "Bretting, Peter" <Peter.Bretting@ars.usda.gov>, "marcelofernandes.oliveira@embr" <marcelofernandes.oliveira@embrapa.br>, Steven Cannon <steven.cannon@ARS.USDA.GOV>, "François Belzile" <François.Belzile@fsaa.ulaval.ca>

날짜 : 2015년 4월 18일(토) 08:21:44

제목 : Follow Up: Glycine meeting Seattle, April 15-17

Folks,

First a big thanks to all that attended– especially those that flew great distances! All of your contribution and inputs are greatly appreciated. I think we had a good meeting and there are many areas of opportunity to work together to better exploit our genebank collections to improve soybean breeding for the future. We identified some first steps where we can work together, but I think we will need to work hard to keep momentum going.

Please let me know if you have any comments or want to make any modifications to the attached report.

Sincerely, Scott Jackson

--

Peter Wenzl

DivSeek Liaison

Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]

www.croptrust.org

Securing our Food, Forever

--
Peter Wenzl
DivSeek Liaison
Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]
www.croptrust.org

Securing our Food, Forever

Bretting, Peter

From: Peter Wenzl <peter.wenzl@croptrust.org>
Sent: Tuesday, May 12, 2015 11:09 AM
To: Scott Allen Jackson
Cc: Cannon, Steven; Suk-Ha Lee; Jones Jennifer; Abberton Michael (IITA); Cannon Steven [AGRON]; Joost Richard; Nelson, Randall - ARS; Okamuro, Jack; 邱丽娟 QIU, Lijuan; 李英慧; Masao Ishimoto; Akito Kaga; Bob Stupar; Ricardo Abdelnoor; Bretting, Peter; marcelofernandes.oliveira@embr; François Belzile
Subject: Re: Follow Up: Glycine meeting Seattle, April 15-17

Dear all,

I've talked to my colleagues here at the Crop Trust who're developing and maintaining the Genesys portal (https://www.genesys-pgr.org/welcome). I hear that it would be quite straightforward to provide room for assembling a global inventory of soybean accessions on this platform, if you wish to do so.

Genesys was designed for the purpose of assembling global inventories of genebank accessions (accession IDs, genebank IDs & passport data). That fact that soybean is not an Annex 1 crop of the International Treaty doesn't matter.

Best regards, Peter

On Wed, May 6, 2015 at 8:40 PM, Scott Allen Jackson <sjackson@uga.edu> wrote:
Thanks for following up Steven,

several issues here:
1) on the funding. I think you are right that initially, at least, individual but coordinated efforts will be necessary. As a group, I think we can provide letters of support stating that the proposed research is part of a larger international initiative.

2) On the collection of accessions (both genotype and non-genotyped). Randy Nelson had 'volunteered' to start the process of collecting all the accessions to begin global inventory. Perhaps, as part of that, people could indicate which ones are genotyped and how. I suspect that much of of the collections have been genotyped in the past, but probably not sequence-based (e.g. SSRs), so the informaiton may not be of much use for what we envisage.

As an aside, we had discussed having an intern work on the data analysis. But the more I think about it, I think this is going to be more of a postdoctoral project—as you indicate below.

scott

> On May 1, 2015, at 12:09 PM, Cannon, Steven <Steven.Cannon@ARS.USDA.GOV> wrote:
>
> Speaking for myself: I am still very enthusiastic about the group's plans, but have been swamped, .
>
> I do, however, have a little progress to report. A student group that I have been advising has worked out a very fast GBS variant-calling and imputation pipeline for soybean. It is processing a lane of GBS data (384 samples multiplexed) in about two hours. This uses something like the "flanking sequencing database"

duplicate email trail removed

Bretting, Peter

From:	Peter Wenzl <peter.wenzl@divseek.org>
Sent:	Friday, May 08, 2015 7:21 PM
To:	Bretting, Peter
Cc:	Susan McCouch; Andreas Graner (IPK); David Marshall; Elizabeth Arnaud (Bioversity); Emily Marden (UBC); Rajeev Varshney (ICRISAT & GCP); Ruaraidh Sackville Hamilton (IRRI); Sarah Ayling (TGAC); Daniele Manzella; Ruth Bastow; Wayne Powell (CGIAR CO)
Subject:	Re: Introduction to DivSeek document
Attachments:	DivSeek - an introduction.docx

Dear all,

Please find attached the Word version of the same document for easier editing.

All the best, Peter

--
Peter Wenzl
DivSeek Liaison
Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]
www.croptrust.org

Securing our Food, Forever

On Fri, May 8, 2015 at 9:00 PM, Bretting, Peter <Peter.Bretting@ars.usda.gov> wrote:

Hi Susan—the document review process you've outlined is fine with me. It's important to impose "version control."

Could you or the JFU provide the text of the "DivSeek—An Introduction" as a Word document (.doc or .docx format)? Track Changes and Comments in Word greatly facilitate addition comments or proposing edits. The Adobe .pdf format can be problematic for that.

Thanks!

Peter

Peter Bretting

USDA/ARS Office of National Programs

Room 4-2212, Mailstop 5139

5601 Sunnyside Avenue

Beltsville, MD 20705-5139

Phone 1.301.504.5541

Fax 1.301.504.6191

Mobile Phone [redacted]

E-mail peter.bretting@ars.usda.gov

Web site: http://www.ars.usda.gov/research/programs/programs.htm?NP_CODE=301

From: [redacted] **On Behalf Of** Susan McCouch
Sent: Monday, May 04, 2015 11:58 AM
To: Andreas Graner (IPK); David Marshall; Elizabeth Arnaud (Bioversity); Emily Marden (UBC); Bretting, Peter; Rajeev Varshney (ICRISAT & GCP); Ruaraidh Sackville Hamilton (IRRI); Sarah Ayling (TGAC); Susan McCouch
Cc: Peter Wenzl; Daniele Manzella; Ruth Bastow; Wayne Powell (CGIAR CO); Susan McCouch
Subject: Introduction to DivSeek document

Dear DivSeek Steering Committee members,

Several of you have asked about the status of the 'Introduction to DivSeek' document discussed during our phone meeting on 28 April, and whether you are free to share it with interested parties.

At this time, that document should be considered a draft. We are interested in your comments and feedback so we can make changes accordingly. Until the document is finalized, we ask that you please do not circulate it.

We look forward to hearing from you about any changes you feel would improve it. The most helpful way to do this would be tp send the document around to SC and JFU members (cc'd above), with tracked changes, or tracked comments indicating specific edits, additions or requests for further clarification. Once we receive your comments, I will work with the JFU to integrate them and we will then bring the document back around for discussion during the in-person SC meeting on May 28 in Rome.

Once the document is finalized, our plan is to make it public over the DivSeek website.

In general, when DivSeek documents are first brought up for discussion in the SC, please treat them as drafts. Once these documents have been vetted through the SC and JFU, and your input has been integrated, my recommendation would be to make them public via the DivSeek website.

Using the DivSeek website to release DivSeek documents to the public will ensure that we have only one version of a document circulating, and that further comments and discussion surrounding that document can be dealt with through information channels established based on the governance structure and principles of DivSeek.

Please let me know if you agree with this recommendation. We are currently working to update and restructure the DivSeek website to improve navigability and make it easier for users to find information.

The SC will have a chance to review the updated web-site at our May 28 meeting in Rome, and your input will be essential as we endeavor to share information with our stakeholders and other interested parties.

Best regards,

Susan

--

Susan McCouch
Professor, Plant Breeding & Genetics
Cornell University
162 Emerson Hall
Ithaca, NY 14853-1901
Phone: +1 607-255-0420
Fax: +1 607-255-6683
Email: srm4@cornell.edu or mccouch@cornell.edu
Alternate Email: [redacted]

DivSeek – An Introduction

By the Joint Facilitation Unit[1]

What is DivSeek?

Over the coming decades, humanity is faced with the challenge of producing sufficient and sufficiently nutritious food for a growing human population in a climate-challenged world. Crop diversity provides the biological foundation for crop improvement and represents a key resource for addressing food security in an environmentally sustainable manner. Coincidently, game-changing technologies such as DNA-sequencing, remote sensing, precision phenotyping, and 'big-data' analytics are enabling a **paradigm shift** in the way natural genetic variation can be investigated and used to develop high-yielding, nutritious, climate-ready crops.

Diversity Seek (DivSeek) is a global initiative motivated by this paradigm shift. The DivSeek initiative builds on other initiatives that originated independently, such as [a] the 3,000 Rice Genomes project at IRRI,[2] and the Seeds of Discovery[3] project at CIMMYT; [b] the "Digital Seedbank" concept envisioned by the Global Plant Council (GPC)[4]; and [c] the Asilomar meeting with community members, supported by a technical study[5], both organized by the Global Crop Diversity Trust (Trust; Figure 1).[6]

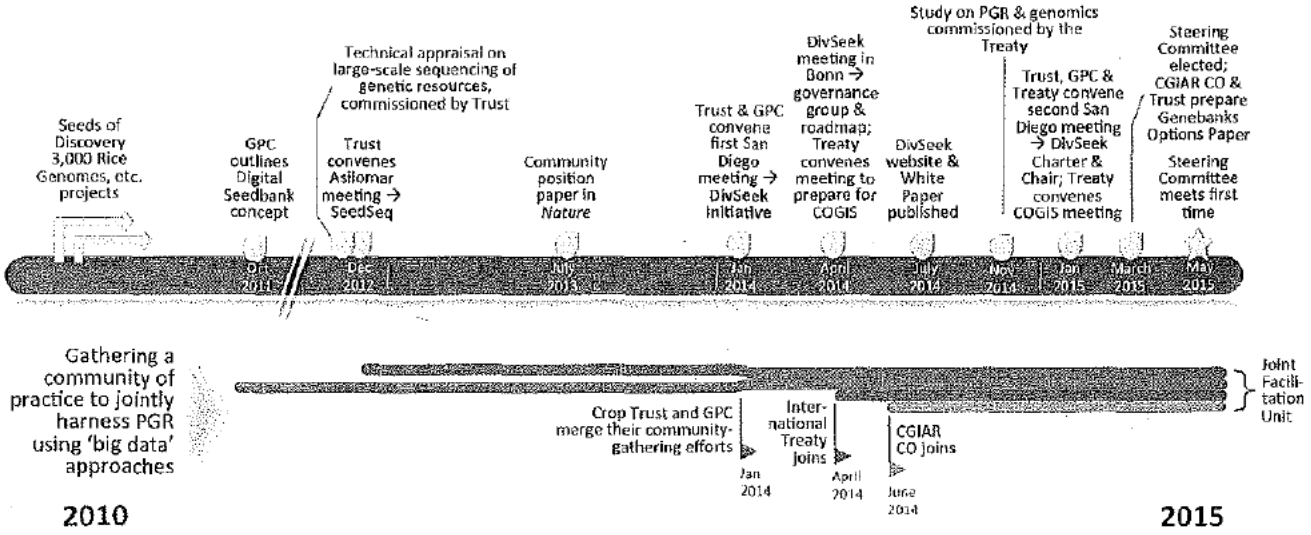


Figure 1: DivSeek timeline.

---

[1] Peter Wenzl, Daniele Manzella, Ruth Bastow, Wayne Powell

[2] The 3,000 Rice Genomes Project [http://www.gigasciencejournal.com/content/pdf/2047-217X-3-7.pdf]

[3] Seeds of Discovery project website [http://seedsofdiscovery.org]

[4] The Global Plant Council: increasing the impact of plant research to meet global challenges [http://www.epsoweb.org/file/1093]

[5] Technical appraisal of strategic approaches to large-scale germplasm evaluation [http://agro.biodiver.se/wp-content/uploads/2012/12/Technical-appraisal-NGS-for-genebanks-please-comment.pdf]

[6] Agriculture: feeding the future [https://www.dropbox.com/s/fekmdmhdbjkss6i/Community%20position%20paper.pdf?dl=0]

During 2013, the Crop Trust and the GPC coordinated their efforts to convene the Plant Genetic Resources (PGR) research community around the idea of using 'big data' approaches to harness PGR to address global food security issues. This process culminated in a jointly organized meeting in San Diego in Jan 2014, which gave rise to a global initiative, referred to as "Diversity Seek", or 'DivSeek'.

In the course of 2014, the Secretariat of the International Treaty on Plant Genetic Resources (Treaty), who were preparing for a consultation on a Global Information System for PGR for Food and Agriculture (PGRFA) and had commissioned a second technical study on PGR and genomics[7], and the CGIAR Consortium Office (CGIAR CO), representing the largest network of international genebanks, both joined the DivSeek initiative.

As an outcome of this process, the DivSeek Joint Facilitation Unit (JFU) comprising one member per each of the four participating organizations was formed. During 2014 the JFU and others began to build the foundations of the DivSeek Initiative with the publication of a White Paper (http://www.divseek.org/white-paper) and the development of a nascent governance structure. In January 2015 the first DivSeek Partner's Assembly was held, a Charter for the initiative was adopted, and the Assembly Chair elected. The DivSeek Steering Committee members were elected in March from amongst the representatives of the DivSeek member organizations that have adopted the Charter (Figure 1).

Purpose

DivSeek brings together genebanks, breeders, crop scientists, and database and computational experts to enhance the use of genetic diversity for crop improvement. The rationale behind the initiative has been outlined previously in a White Paper (http://www.divseek.org/white-paper). The DivSeek Charter, adopted by a large number or organizations from around the world at the first DivSeek Assembly states:[8]

> *"The mission of DivSeek is to cross-link, support and add value to individual activities that harness the power of crop diversity for food and nutritional security and societal and economic benefits, by enabling breeders and researchers to mobilize genetic variation in order to accelerate crop improvement."*

Principles

"If you want to go fast, go alone. If you want to go far, go together". This African proverb describes one of the central tenets of DivSeek. Another core element of DivSeek is *diversity* itself, as a driver not only for genetic progress in crop improvement but also to enable innovation in science.

In the preparatory meetings leading up to the 2015 Partners' Assembly, DivSeek has been compared to a **comb** that links like-minded partners to create synergies among, and add value to individual projects, harnessing genetic resources in a variety of crops. The teeth of the comb

---

[7] Plant genetic resources and genomics: mainstreaming agricultural research through genomics [http://www.planttreaty.org/sites/default/files/ITPGRFA_BS005e.pdf]

[8] DivSeek Charter [https://www.dropbox.com/s/aj9qcsa89pvyveh/DivSeek%20Charter.pdf?dl=0]

represent individual projects characterizing genetic resources of different crops. The back of the comb represents DivSeek acting as cross-crop platform that connects individual projects in areas of shared interests and common challenges and enables them to leverage each others strengths to move forward in a synchronous and coordinated way (Figure 2, left half).

An alternative image to visualize the 'DivSeek approach' would be that of a **multi-colored umbrella**, which unfolds its diversity of colors (individual crop-specific projects) through the synchronous movement of its frame pieces (common DivSeek activities) that support and expand its colorful canopy (Figure 2, right half).

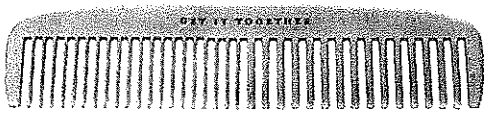


Figure 2: DivSeek in two images.

Other key elements of the DivSeek initiative, discussed on several occasions during past meetings and outlined in the DivSeek Charter, include:

- **A community-driven spirit based on voluntary cooperation:** The DivSeek community has come together as a result of research-community members interested in plant genetic resources (PGR) seeking to synchronize and mutually leverage their research agendas.
- **The science-driven agenda on which DivSeek is based:** DivSeek aims to accelerate innovation in the way genetic resources are utilized by linking a growing body of genotypic, phenotypic and eco-geographic characterization data to physical germplasm accessions and associated passport data.
- **A modular approach to information management:** DivSeek advocates the use of data standards to enable information integration and interoperability among data sets and repositories. This includes the use of unique germplasm identifiers to link DivSeek data to a global information system for plant genetic resources.
- **A common 'rights management' framework for sharing PGR-associated data:** DivSeek will work on a broadly acceptable framework that stimulates appropriate sharing and use of PGR characterization/evaluation data in an open-data framework.
- **The importance of capacity-building efforts:** DivSeek aims to offer cross-crop learning and capacity-development opportunities to provide broad access to enabling tools and platforms for the analysis of genetic resources.

Intended Impacts

Based on the DivSeek Mission statement, **demand-driven agricultural research**, and crop improvement in particular, are the principal targets and beneficiaries of DivSeek. DivSeek will

advocate for a broadening of the genetic base available to breeding programs, thus diversifying their options for responding to the challenges ahead.

DivSeek will also feed into fundamental **discovery-driven plant research**, thus stimulating discoveries that will create spillover effects into crop improvement and sustainable crop production.

Importantly, DivSeek will enable **genebanks** to fundamentally transform themselves into research centers that proactively investigate the genetic potential of their holdings. This transformation is expected to amplify impacts in both the demand and discovery-driven research domains.

Summary of 2015 events

DivSeek Assembly (January)

The first DivSeek Partners' Assembly in January discussed and adopted a **Charter for DivSeek** and appointed Susan McCouch as the Assembly's Chairperson.[9] To this date, the Charter has been endorsed by 57 of the 69 organizations that originally voiced their interest in joining DivSeek.[10] Twelve organizations based in Australia, Colombia, Côte d'Ivoire, France, Indonesia, Italy, Kenya, New Caledonia, Thailand and the UK, who previously had expressed their interest in joining DivSeek, have not (yet) endorsed the Charter.

Several ongoing efforts to systematically **mine genetic resources** were presented, including the Seeds of Discovery project on maize and wheat genetic resources at CIMMYT, the 3,000 Rice Genomes Project at IRRI, the BRIDGE project on barley at IPK and the 3,000 Chickpea Genomes project and related efforts on other crops at ICRISAT.

A second series of presentations covered **information-management**-related topics. These included the Planteome ontology platform lead by Oregon State University, the Germinate repository for genetic, phenotypic and environmental data at the James Hutton Institute, the high-performance cyber-infrastructure at iPlant, the GOBII initiative on deploying high-density marker data in crop improvement led by Cornell University, and the ELIXIR genomics platform at EMBL-EBI. It would appear that each of these platforms has unique features that could cover certain aspects of information-management requirements of DivSeek participants working on a variety of crops. The challenge may be to achieve interoperability among these platforms, so that they can be combined and act as 'building blocks' in support of individual DivSeek projects.

Researchers from Arizona State University (ASU) presented a project proposal that seeks to explore **governance options** for DivSeek by documenting lessons learned by other research initiatives dealing with genetics and genetic resources. Presentations outlining national-genebanks', the CGIAR's, and funder organizations' perspectives on DivSeek complemented the overall picture and outlined priorities and areas of opportunity.

---

[9] Report on the governance session of the 2015 DivSeek Assembly https://www.dropbox.com/s/pk32130bg5tx067/DivSeek%20Assembly%20report%20-%20governance%20session.pdf?dl=0

[10] DivSeek Charter [https://www.dropbox.com/s/aj9qcsa89pvyveh/DivSeek%20Charter.pdf?dl=0]

The feedback received to a survey distributed during the Assembly validated the overarching importance of information management for the majority of respondents. Fifty percent of all answers were related to data standards, data repositories, data-analysis tools, or assistance with data management. Topics related to community building & networking (19%) and research strategies (15%) were also mentioned frequently. The remaining 16% of answers had to do with PGR rights management, funding for individual projects, datasets for individual crops and capacity building.[11]

Expert Consultation on the Global Information System (COGIS; January)

In the context of the ongoing COGIS process coordinated by the International Treaty on Plant Genetic Resources for Food and Agriculture, a meeting was held during which a vision paper on a Global Information System for PGRFA was presented[12] and the use of Permanent Unique Identifiers (PUID) for tracking germplasm was discussed.[13] A task force subsequently concluded that Digital Object Identifiers (DOI) were the preferable system to track PGRFA and their basic passport data.[14] Interestingly and importantly, the same DOI system has already been implemented by the Global Biodiversity Information Facility to track biodiversity datasets. Also, the 3,000 Rice Genomes Project already uses DOI to identify primary data sets.[2]

Strengthening of the Joint Facilitation Unit (January)

As of January 2015, the JFU had relied on the part-time engagement of four staff from the four organizations participating in the JFU to implement DivSeek (Hannes Dempewolf, Daniele Manzella, Ruth Bastow, Wayne Powell). In anticipation of an accelerating DivSeek agenda during 2015, the Crop Trust hired Peter Wenzl, who previously led the Seeds of Discovery project at CIMMYT, considered to be one of the 'blueprint projects' for DivSeek to replace Hannes as the Crop Trust representative on the JFU. Peter is now working close-to full-time on DivSeek matters. The Secretariat of the International Treaty also strengthened its contribution to JFU discussions with contributions from Francisco Lopez on matters related to the Global Information System on PGRFA. Other staff from the four organizations may be called on to contribute to DivSeek activities on a case-by-case basis, particularly on matters such as communication and fund raising.

---

[11] Report on the technical session of the 2015 DivSeek Assembly https://www.dropbox.com/s/8eebwobfuiv6w6j/DivSeek%20Assembly%20report%20-%20technical%20session.pdf?dl=0

[12] The Vision Paper for the Global Information System on Plant Genetic Resources for Food and Agriculture [http://www.planttreaty.org/sites/default/files/cogis1w5.pdf]

[13] First meeting of the expert consultation on the global information system on plant genetic resources for food and agriculture [http://www.planttreaty.org/sites/default/files/COGIS1re.pdf]

[14] Summary report of the task force on permanent global and unique identifiers for PGRFA in the context of the Global Information System of Article 17 [https://www.dropbox.com/s/7qroz110yrppnxi/PUID%20task%20force%20recommendations%20.pdf?dl=0]

Election of the Steering Committee (March)

Following the Assembly, eight Steering Committee members were elected through electronic voting out of a pool of 16 candidates nominated from and by the DivSeek Member Organizations. Members of the Steering Committee cover a broad spectrum of skills, ranging from genetic-resources & genebank management, genomics, bioinformatics, ontologies and governance. Four of the nine Steering Committee members (including the Chair of the Assembly) are women. It was noted that representatives from developing countries and plant breeding professionals are underrepresented at this time.

Genebanks Options Paper (March-April)

Staff of the CGIAR Consortium Office and the Crop Trust, and genebank managers from CGIAR centers jointly drafted a document that outlines options for funding genebank operations beyond the conclusion of the current Genebanks CRP. In addition to discussing funding options for basic genebank operations, the paper contains a chapter that highlights opportunities for mining genebanks using 'big data'-driven approaches. Chapter 5 refers to DivSeek and highlights the CGIAR centers' genebanks' prominent role in the initiative.[15]

New Projects in the 'DivSeek Domain' (March-April)

It has been encouraging and exiting to see that new work on a variety of DivSeek-related objectives has already started with support from different organizations and research communities. This may be a sign that the time is ripe for DivSeek to gain traction. Examples of new activities and projects include:

- A research project at Arizona State University (ASU) to explore governance options for DivSeek based on lessons learned in similar initiatives, to be co-funded by the International Treaty and the Crop Trust. We expect that this project will outline empirically validated criteria associated with success of community-driven initiatives in genetics and genetic resources.
- A Benefit Sharing Fund-sponsored project to build and test a platform for allocating PUID in rice, led by the Indonesian Agency of Agriculture Research and Development (IAARD). The PUID and the application-programming interface (API) to be developed are designed to enable linkages between rice DivSeek data and the multi-lateral system of the International Treaty.
- The Crop Wild Relatives Project is developing software and data visualization tools and populating the Germinate data repository, developed by the James Hutton Institute (JHI), with new genotypic and phenotypic data related to the use of wild relatives of sunflower and rice improvement. The same platform has already been implemented by CIMMYT's Seeds of Discovery project to disseminate maize and wheat data, and previously has also been used for barley and potato data.

[15] CGIAR genebanks options paper for FC13 [https://www.dropbox.com/s/57pse1y2b53z7b1/Genebanks%20options%20paper.pdf?dl=0]

- Scott Jackson from the University of Georgia recently convened a global group of researchers interested in soybean genetic resources to elaborate a strategy for soybeans as part of DivSeek.
- Discussions with Roberto Papa from Università Politecnica delle Marche are being initiated to link a recently funded ERA CAP project on common beans with DivSeek. The project includes the genotyping of 10,000 globally sourced accessions.

How will DivSeek be Governed and Managed?

The guidelines for a community-driven DivSeek governance framework have been spelled out in the DivSeek Charter adopted at the Assembly in January 2015. The framework is composed of three governing bodies[16]

- The Partners' Assembly (**PA** for short) is the ultimate decision body and as such determines the strategy of DivSeek,
- The Steering Committee (**SC** for short) guides the implementation of the DivSeek strategy, and
- The Joint Facilitation Unit (**JFU** for short) serves both the **PA** and the **SC**. A draft operational plan will be submitted to the **SC** and the **PA** to describe the day-to-day functions of the JFU.

Setting a strategic direction for DivSeek and guiding its implementation (PA, SC)

The annual PA in its role as a 'DivSeek parliament' determines the strategic direction of DivSeek. It considers and approves work plans, budgets, resource-mobilization plans and annual reports. It also elects an Assembly Chairperson and the members of the SC.

The SC comprises eight members and the Chair of the PA. It meets at least twice a year to oversee and guide the implementation of the DivSeek strategy. It advises the PA about the strategic direction of DivSeek activities and provides input and further develops draft work plans, budgets, resource-mobilization plans and annual reports prepared by the JFU, and endorses final versions of these documents. It also collects and reports to the PA information about interactions among partners, and works with the JFU to prepare DivSeek updates for the constituencies and governing bodies of the four organizations contributing to the JFU.

Implementing DivSeek on a day-to-day basis (JFU)

The JFU operationalizes work plans and facilitates their implementation on a day-to-day basis. It develops drafts work plans, draft budgets, resource-mobilization plans and annual reports for consideration by the SC and the PA. It supports the definition of operational guidelines for implementing DivSeek; mobilizes resources for DivSeek work plans; and administers the JFU budget. It also organizes PA and SC meetings; deals with DivSeek membership requests and recruitment; promotes linkages to other initiatives of relevance for DivSeek; and assists the SC in collecting information about interactions among partners.

[16] DivSeek Charter [https://www.dropbox.com/s/aj9qcsa89pvyveh/DivSeek%20Charter.pdf]

Four organizations with global reach and complementary constituencies contribute one representative each to the JFU to jointly implement DivSeek: the *Global Crop Diversity Trust* (Crop Trust), the *Global Plant Council* (GPC), the *Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture* (International Treaty), and the *CGIAR Consortium Office* (CGIAR-CO) (Figure 6).

Each of these four organizations contributes a wide range of expertise and technical knowledge to the DivSeek JFU, and examples of some of these are provided in the diagram below.

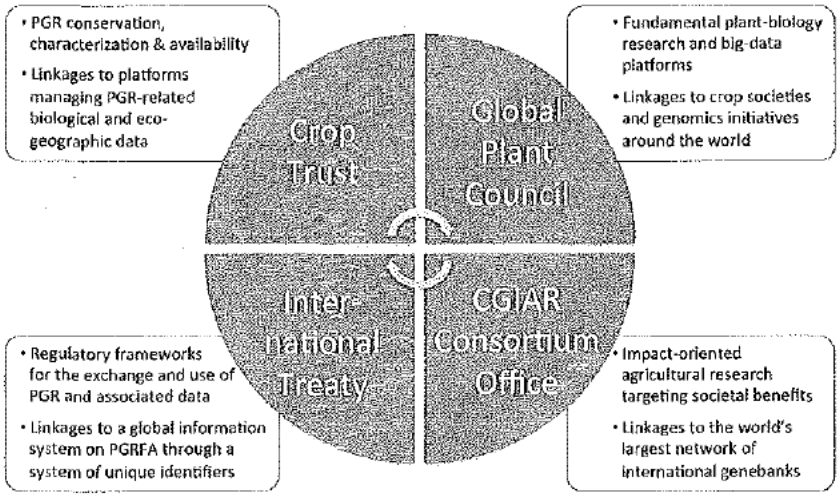


Figure 3: A non-exhaustive set of examples of how the Crop Trust, GPC, International Treaty and CGIAR-CO contribute to the Joint Facilitation Unit

Towards a DivSeek Program of Work

One of the first tasks of the Steering Committee will be the development of a program of work. To assist the SC, the JFU provide appropriate materials for elaborating a Program of Work, for example the JFU will prepare a 'menu' of suggested areas of DivSeek activities. Initially this will be based on the priorities identified by partners during the 2015 Assembly (outlined below), the priorities will compared to the goals of existing projects of relevance for DivSeek, which will be identified in a preliminary landscaping study (briefly outlined below) in advance of the Steering Committee meeting in May. This approach should allow us to suggest a portfolio of pilot projects that could provide 'building blocks' for DivSeek.

A second component of an initial Program of Work will deal with DivSeek management and other 'housekeeping' issues. This component builds upon the DivSeek governance framework. It will likely include activities such as:

- Convening Partner Assembly and Steering Committee meetings
- A communication strategy
- Building linkages with other relevant initiatives
- Refining operational guidelines for DivSeek governance bodies, for example, based on insights provided by the ASU research project

Priority Areas Identified by Partners

The survey conducted as part of the 2015 Partners' Assembly[17], as well as recent insights from discussions about implementing DivSeek in selected crop communities, point to the following priority areas of work:

- Information management
  - Data standards
    - Permanent Unique Identifiers (PUID) for germplasm and datasets
    - (Meta)data standards for high-density molecular data (GbS, SNP arrays)
    - (Meta)data standards for various types of phenotypic data
  - Software platforms
    - Data repositories and tools for distributing, processing and visualizing validated datasets
    - High-performance computation platforms and repositories for storing and processing primary data
    - Data bases and tools for gathering and managing primary data
- Community building & networking
  - Inventory of research approaches and lessons learned
  - Community-building: social network of PGR researchers
  - DivSeek as a neutral platform to broker germplasm and data exchange among organizations working on a particular crop
  - Joint advocacy efforts: public awareness and fundraising
- Research approaches & platforms
  - Accession-sampling strategies (inbred, outcrossing crops, etc.)
  - Genotyping platforms
  - Phenotyping methods
  - Targeting and quantifying intended impacts (crop improvement programs, upstream research, genebank operations, etc.)
- Rights management
  - Measures to incentivize sharing of germplasm for DivSeek research for PGR regulated under the International Treaty and the Convention on Biological Diversity
  - A commonly accepted open-data rights framework for sharing DivSeek data within and across projects
  - The role of public-private partnerships in DivSeek
- Capacity building
  - Training in data management, curation and analysis
  - Identification of other priority areas and targeted audiences for capacity building efforts

[17] Report on the technical session of the 2015 DivSeek Assembly https://www.dropbox.com/s/8eebwobfuiv6w6j/DivSeek%20Assembly%20report%20-%20technical%20session.pdf?dl=0

Landscape of Ongoing Projects

The JFU will conduct a preliminary landscaping study before the SC meeting in May. This study will take into account and expand upon previous efforts to identify 'DivSeek'-like projects across a variety of crops. We will attempt to identify a 'core collection' of ongoing projects by maximizing the 'diversity' according to three sets of criteria that can be used to describe individual projects:

- The importance of a particular crop for agriculture
- Factors that shape research approaches
- The regulatory framework(s) applicable to a crop and its wild relatives

Each of these three categories contains subcategories designed to broadly capture features of relevance to categorize different crops for the purposes of DivSeek (Figure 4).

The second area the landscape study will focus on the information-management domain. Data standards, data repositories and high-performance computational environments for data processing may be the three principal areas to investigate.

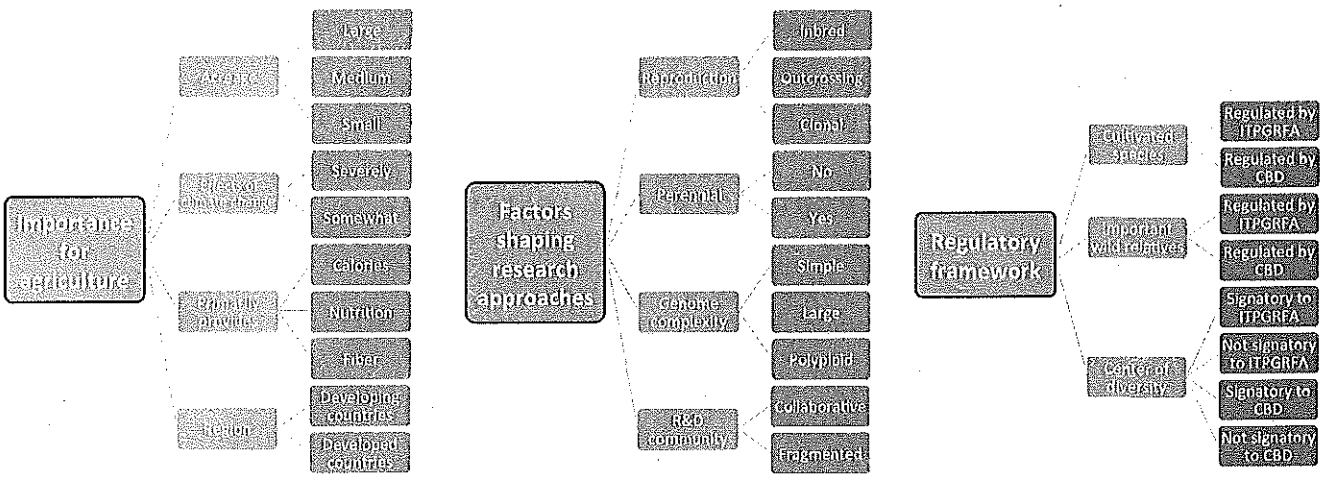


Figure 4: Possible set of criteria for choosing a subset of DivSeek-like projects across a variety of crops.

Bretting, Peter

From: E. Marden [redacted]
Sent: Tuesday, May 05, 2015 6:59 PM
To: Susan McCouch
Cc: Andreas Graner (IPK); David Marshall; Elizabeth Arnaud (Bioversity); Bretting, Peter; Rajeev Varshney (ICRISAT & GCP); Ruaraidh Sackville Hamilton (IRRI); Sarah Ayling (TGAC); Peter Wenzl; Daniele Manzella; Ruth Bastow; Wayne Powell (CGIAR CO); Susan McCouch
Subject: Re: Introduction to DivSeek document
Attachments: DivSeek - an introduction - EM comments.pdf

Hi Susan,

I had a number of comments on the document. I believe I only had the PDF version and so have incorporated them in comment bubbles. There are other issues that are not addressed in the document - such as what is a Program of Work, and how does DivSeek interact with it? Is it simply a voluntary affiliation, like the DivSeek types of projects mentioned in the report? Or is there some kind of 'approval' and association (and potentially funding)? If a project is a DivSeek project, must they agree to principles of sharing? I admit this is a difficult question as those principles are not yet articulated. For the moment, it may be helpful to include something as a placeholder as I believe others will have the same questions about how DivSeek works.

In addition, I raised a question about COGIS and the relationship to DivSeek. I think it is important to state whether there is a relationship or whether it is just a related but independent effort.

Other comments should be self-explanatory.

Best regards,

Emily

On May 4, 2015, at 8:58 AM, Susan McCouch <srm4@cornell.edu> wrote:

> Dear DivSeek Steering Committee members,
>
> Several of you have asked about the status of the 'Introduction to DivSeek' document discussed during our phone meeting on 28 April, and whether you are free to share it with interested parties.
>
> At this time, that document should be considered a draft. We are interested in your comments and feedback so we can make changes accordingly. Until the document is finalized, we ask that you please do not circulate it.
>
> We look forward to hearing from you about any changes you feel would improve it. The most

duplicate email trail removed

DivSeek – An Introduction

By the Joint Facilitation Unit[1]

What is DivSeek?

Over the coming decades, humanity is faced with the challenge of producing sufficient and sufficiently nutritious food for a growing human population in a climate-challenged world. Crop diversity provides the biological foundation for crop improvement and represents a key resource for addressing food security in an environmentally sustainable manner. Coincidently, game-changing technologies such as DNA-sequencing, remote sensing, precision phenotyping, and 'big-data' analytics are enabling a **paradigm shift** in the way natural genetic variation can be investigated and used to develop high-yielding, nutritious, climate-ready crops.

Diversity Seek (DivSeek) is a global initiative motivated by this paradigm shift. The DivSeek initiative builds on other initiatives that originated independently, such as [a] the 3,000 Rice Genomes project at IRRI,[2] and the Seeds of Discovery[3] project at CIMMYT; [b] the "Digital Seedbank" concept envisioned by the Global Plant Council (GPC)[4]; and [c] the Asilomar meeting with community members, supported by a technical study[5], both organized by the Global Crop Diversity Trust (Trust; Figure 1).[6]

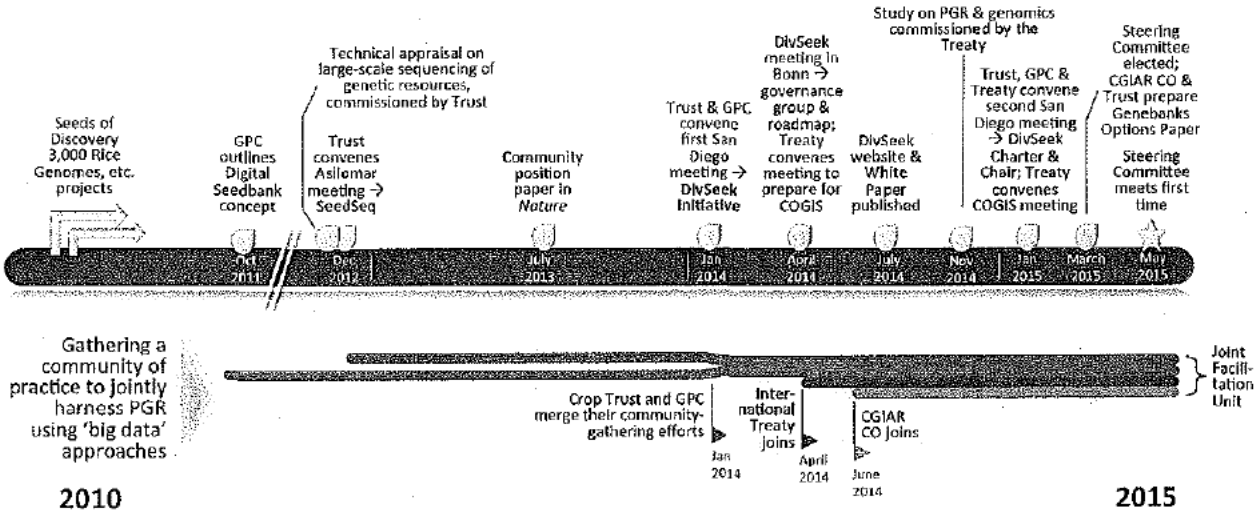


Figure 1: DivSeek timeline.

---

[1] Peter Wenzl, Daniele Manzella, Ruth Bastow, Wayne Powell

[2] The 3,000 Rice Genomes Project [http://www.gigasciencejournal.com/content/pdf/2047-217X-3-7.pdf]

[3] Seeds of Discovery project website [http://seedsofdiscovery.org]

[4] The Global Plant Council: increasing the impact of plant research to meet global challenges [http://www.epsoweb.org/file/1093]

[5] Technical appraisal of strategic approaches to large-scale germplasm evaluation [http://agro.biodiver.se/wp-content/uploads/2012/12/Technical-appraisal-NGS-for-genebanks-please-comment.pdf]

[6] Agriculture: feeding the future [https://www.dropbox.com/s/fekmdmhdbjkss6i/Community%20position%20paper.pdf?dl=0]

During 2013, the Crop Trust and the GPC coordinated their efforts to convene the Plant Genetic Resources (PGR) research community around the idea of using 'big data' approaches to harness PGR to address global food security issues. This process culminated in a jointly organized meeting in San Diego in Jan 2014, which gave rise to a global initiative, referred to as "Diversity Seek", or 'DivSeek'.

In the course of 2014, the Secretariat of the International Treaty on Plant Genetic Resources (Treaty), who were preparing for a consultation on a Global Information System for PGR for Food and Agriculture (PGRFA) and had commissioned a second technical study on PGR and genomics[7], and the CGIAR Consortium Office (CGIAR CO), representing the largest network of international genebanks, both joined the DivSeek initiative.

As an outcome of this process, the DivSeek Joint Facilitation Unit (JFU) comprising one member per each of the four participating organizations was formed. During 2014 the JFU and others began to build the foundations of the DivSeek Initiative with the publication of a White Paper (http://www.divseek.org/white-paper) and the development of a nascent governance structure. In January 2015 the first DivSeek Partner's Assembly was held, a Charter for the initiative was adopted, and the Assembly Chair elected. The DivSeek Steering Committee members were elected in March from amongst the representatives of the DivSeek member organizations that have adopted the Charter (Figure 1).

Purpose

DivSeek brings together genebanks, breeders, crop scientists, and database and computational experts to enhance the use of genetic diversity for crop improvement. The rationale behind the initiative has been outlined previously in a White Paper (http://www.divseek.org/white-paper). The DivSeek Charter, adopted by a large number or organizations from around the world at the first DivSeek Assembly states:[8]

> *"The mission of DivSeek is to cross-link, support and add value to individual activities that harness the power of crop diversity for food and nutritional security and societal and economic benefits, by enabling breeders and researchers to mobilize genetic variation in order to accelerate crop improvement."*

Principles

"If you want to go fast, go alone. If you want to go far, go together". This African proverb describes one of the central tenets of DivSeek. Another core element of DivSeek is *diversity* itself, as a driver not only for genetic progress in crop improvement but also to enable innovation in science.

In the preparatory meetings leading up to the 2015 Partners' Assembly, DivSeek has been compared to a **comb** that links like-minded partners to create synergies among, and add value to individual projects, harnessing genetic resources in a variety of crops. The teeth of the comb

---

[7] Plant genetic resources and genomics: mainstreaming agricultural research through genomics [http://www.planttreaty.org/sites/default/files/ITPGRFA_BS005e.pdf]

[8] DivSeek Charter [https://www.dropbox.com/s/aj9qcsa89pvyveh/DivSeek%20Charter.pdf?dl=0]

represent individual projects characterizing genetic resources of different crops. The back of the comb represents DivSeek acting as cross-crop platform that connects individual projects in areas of shared interests and common challenges and enables them to leverage each others strengths to move forward in a synchronous and coordinated way (Figure 2, left half).

An alternative image to visualize the 'DivSeek approach' would be that of a **multi-colored umbrella**, which unfolds its diversity of colors (individual crop-specific projects) through the synchronous movement of its frame pieces (common DivSeek activities) that support and expand its colorful canopy (Figure 2, right half).

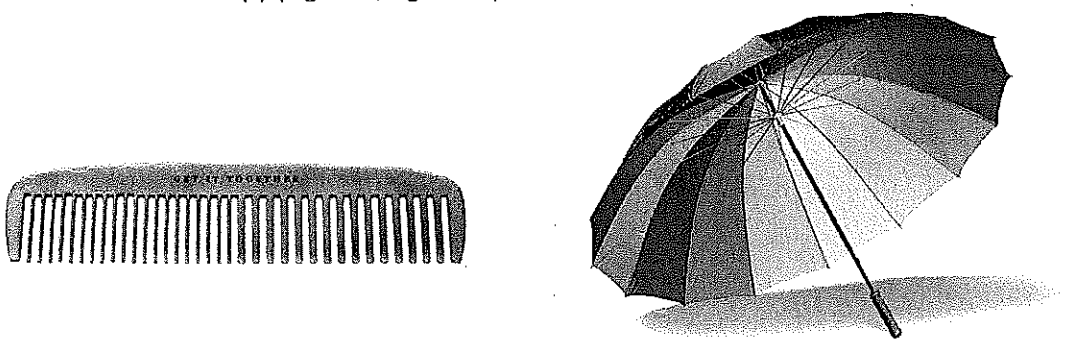


Figure 2: DivSeek in two images.

Other key elements of the DivSeek initiative, discussed on several occasions during past meetings and outlined in the DivSeek Charter, include:

- **A community-driven spirit based on voluntary cooperation:** The DivSeek community has come together as a result of research-community members interested in plant genetic resources (PGR) seeking to synchronize and mutually leverage their research agendas.
- **The science-driven agenda on which DivSeek is based:** DivSeek aims to accelerate innovation in the way genetic resources are utilized by linking a growing body of genotypic, phenotypic and eco-geographic characterization data to physical germplasm accessions and associated passport data.
- **A modular approach to information management:** DivSeek advocates the use of data standards to enable information integration and interoperability among data sets and repositories. This includes the use of unique germplasm identifiers to link DivSeek data to a global information system for plant genetic resources.
- **A common 'rights management' framework for sharing PGR-associated data:** DivSeek will work on a broadly acceptable framework that stimulates appropriate sharing and use of PGR characterization/evaluation data in an open-data framework.
- **The importance of capacity-building efforts:** DivSeek aims to offer cross-crop learning and capacity-development opportunities to provide broad access to enabling tools and platforms for the analysis of genetic resources.

Intended Impacts

Based on the DivSeek Mission statement, **demand-driven agricultural research**, and crop improvement in particular, are the principal targets and beneficiaries of DivSeek. DivSeek will

advocate for a broadening of the genetic base available to breeding programs, thus diversifying their options for responding to the challenges ahead.

DivSeek will also feed into fundamental **discovery-driven plant research**, thus stimulating discoveries that will create spillover effects into crop improvement and sustainable crop production.

Importantly, DivSeek will enable **genebanks** to fundamentally transform themselves into research centers that proactively investigate the genetic potential of their holdings. This transformation is expected to amplify impacts in both the demand and discovery-driven research domains.

Summary of 2015 events

DivSeek Assembly (January)

The first DivSeek Partners' Assembly in January discussed and adopted a **Charter for DivSeek** and appointed Susan McCouch as the Assembly's Chairperson.[9] To this date, the Charter has been endorsed by 57 of the 69 organizations that originally voiced their interest in joining DivSeek.[10] Twelve organizations based in Australia, Colombia, Côte d'Ivoire, France, Indonesia, Italy, Kenya, New Caledonia, Thailand and the UK, who previously had expressed their interest in joining DivSeek, have not (yet) endorsed the Charter.

Several ongoing efforts to systematically **mine genetic resources** were presented, including the Seeds of Discovery project on maize and wheat genetic resources at CIMMYT, the 3,000 Rice Genomes Project at IRRI, the BRIDGE project on barley at IPK and the 3,000 Chickpea Genomes project and related efforts on other crops at ICRISAT.

A second series of presentations covered **information-management**-related topics. These included the Planteome ontology platform lead by Oregon State University, the Germinate repository for genetic, phenotypic and environmental data at the James Hutton Institute, the high-performance cyber-infrastructure at iPlant, the GOBII initiative on deploying high-density marker data in crop improvement led by Cornell University, and the ELIXIR genomics platform at EMBL-EBI. It would appear that each of these platforms has unique features that could cover certain aspects of information-management requirements of DivSeek participants working on a variety of crops. The challenge may be to achieve interoperability among these platforms, so that they can be combined and act as 'building blocks' in support of individual DivSeek projects.

Researchers from Arizona State University (ASU) presented a project proposal that seeks to explore **governance options** for DivSeek by documenting lessons learned by other research initiatives dealing with genetics and genetic resources. Presentations outlining national-genebanks', the CGIAR's, and funder organizations' perspectives on DivSeek complemented the overall picture and outlined priorities and areas of opportunity.

---

[9] Report on the governance session of the 2015 DivSeek Assembly https://www.dropbox.com/s/pk32130bg5tx067/DivSeek%20Assembly%20report%20-%20governance%20session.pdf?dl=0

[10] DivSeek Charter [https://www.dropbox.com/s/aj9qcsa89pvyveh/DivSeek%20Charter.pdf?dl=0]

The feedback received to a survey distributed during the Assembly validated the overarching importance of information management for the majority of respondents. Fifty percent of all answers were related to data standards, data repositories, data-analysis tools, or assistance with data management. Topics related to community building & networking (19%) and research strategies (15%) were also mentioned frequently. The remaining 16% of answers had to do with PGR rights management, funding for individual projects, datasets for individual crops and capacity building.[11]

Expert Consultation on the Global Information System (COGIS; January)

In the context of the ongoing COGIS process coordinated by the International Treaty on Plant Genetic Resources for Food and Agriculture, a meeting was held during which a vision paper on a Global Information System for PGRFA was presented[12] and the use of Permanent Unique Identifiers (PUID) for tracking germplasm was discussed.[13] A task force subsequently concluded that Digital Object Identifiers (DOI) were the preferable system to track PGRFA and their basic passport data.[14] Interestingly and importantly, the same DOI system has already been implemented by the Global Biodiversity Information Facility to track biodiversity datasets. Also, the 3,000 Rice Genomes Project already uses DOI to identify primary data sets.[2]

Strengthening of the Joint Facilitation Unit (January)

As of January 2015, the JFU had relied on the part-time engagement of four staff from the four organizations participating in the JFU to implement DivSeek (Hannes Dempewolf, Daniele Manzella, Ruth Bastow, Wayne Powell). In anticipation of an accelerating DivSeek agenda during 2015, the Crop Trust hired Peter Wenzl, who previously led the Seeds of Discovery project at CIMMYT, considered to be one of the 'blueprint projects' for DivSeek to replace Hannes as the Crop Trust representative on the JFU. Peter is now working close-to full-time on DivSeek matters. The Secretariat of the International Treaty also strengthened its contribution to JFU discussions with contributions from Francisco Lopez on matters related to the Global Information System on PGRFA. Other staff from the four organizations may be called on to contribute to DivSeek activities on a case-by-case basis, particularly on matters such as communication and fund raising.

Election of the Steering Committee (March)

---

[11] Report on the technical session of the 2015 DivSeek Assembly https://www.dropbox.com/s/8eebwobfuiv6w6j/DivSeek%20Assembly%20report%20-%20technical%20session.pdf?dl=0

[12] The Vision Paper for the Global Information System on Plant Genetic Resources for Food and Agriculture [http://www.planttreaty.org/sites/default/files/cogis1w5.pdf]

[13] First meeting of the expert consultation on the global information system on plant genetic resources for food and agriculture [http://www.planttreaty.org/sites/default/files/COGIS1re.pdf]

[14] Summary report of the task force on permanent global and unique identifiers for PGRFA in the context of the Global Information System of Article 17 [https://www.dropbox.com/s/7qroz110yrppnxi/PUID%20task%20force%20recommendations%20.pdf?dl=0]

Following the Assembly, eight Steering Committee members were elected through electronic voting out of a pool of 16 candidates nominated from and by the DivSeek Member Organizations. Members of the Steering Committee cover a broad spectrum of skills, ranging from genetic-resources & genebank management, genomics, bioinformatics, ontologies and governance. Four of the nine Steering Committee members (including the Chair of the Assembly) are women. It was noted that representatives from developing countries and plant breeding professionals are underrepresented at this time.

Genebanks Options Paper (March-April)

Staff of the CGIAR Consortium Office and the Crop Trust, and genebank managers from CGIAR centers jointly drafted a document that outlines options for funding genebank operations beyond the conclusion of the current Genebanks CRP. In addition to discussing funding options for basic genebank operations, the paper contains a chapter that highlights opportunities for mining genebanks using 'big data'-driven approaches. Chapter 5 refers to DivSeek and highlights the CGIAR centers' genebanks' prominent role in the initiative.[15]

New Projects in the 'DivSeek Domain' (March-April)

It has been encouraging and exiting to see that new work on a variety of DivSeek-related objectives has already started with support from different organizations and research communities. This may be a sign that the time is ripe for DivSeek to gain traction. Examples of new activities and projects include:

- A research project at Arizona State University (ASU) to explore governance options for DivSeek based on lessons learned in similar initiatives, to be co-funded by the International Treaty and the Crop Trust. We expect that this project will outline empirically validated criteria associated with success of community-driven initiatives in genetics and genetic resources.
- A Benefit Sharing Fund-sponsored project to build and test a platform for allocating PUID in rice, led by the Indonesian Agency of Agriculture Research and Development (IAARD). The PUID and the application-programming interface (API) to be developed are designed to enable linkages between rice DivSeek data and the multi-lateral system of the International Treaty.
- The Crop Wild Relatives Project is developing software and data visualization tools and populating the Germinate data repository, developed by the James Hutton Institute (JHI), with new genotypic and phenotypic data related to the use of wild relatives of sunflower and rice improvement. The same platform has already been implemented by CIMMYT's Seeds of Discovery project to disseminate maize and wheat data, and previously has also been used for barley and potato data.

[15] CGIAR genebanks options paper for FC13 [https://www.dropbox.com/s/57pse1y2b53z7b1/Genebanks%20options%20paper.pdf?dl=0]

- Scott Jackson from the University of Georgia recently convened a global group of researchers interested in soybean genetic resources to elaborate a strategy for soybeans as part of DivSeek.
- Discussions with Roberto Papa from Università Politecnica delle Marche are being initiated to link a recently funded ERA CAP project on common beans with DivSeek. The project includes the genotyping of 10,000 globally sourced accessions.

How will DivSeek be Governed and Managed?

The guidelines for a community-driven DivSeek governance framework have been spelled out in the DivSeek Charter adopted at the Assembly in January 2015. The framework is composed of three governing bodies[16]

- The Partners' Assembly (**PA** for short) is the ultimate decision body and as such determines the strategy of DivSeek,
- The Steering Committee (**SC** for short) guides the implementation of the DivSeek strategy, and
- The Joint Facilitation Unit (**JFU** for short) serves both the **PA** and the **SC**. A draft operational plan will be submitted to the **SC** and the **PA** to describe the day-to-day functions of the JFU.

Setting a strategic direction for DivSeek and guiding its implementation (PA, SC)

The annual PA in its role as a 'DivSeek parliament' determines the strategic direction of DivSeek. It considers and approves work plans, budgets, resource-mobilization plans and annual reports. It also elects an Assembly Chairperson and the members of the SC.

The SC comprises eight members and the Chair of the PA. It meets at least twice a year to oversee and guide the implementation of the DivSeek strategy. It advises the PA about the strategic direction of DivSeek activities and provides input and further develops draft work plans, budgets, resource-mobilization plans and annual reports prepared by the JFU, and endorses final versions of these documents. It also collects and reports to the PA information about interactions among partners, and works with the JFU to prepare DivSeek updates for the constituencies and governing bodies of the four organizations contributing to the JFU.

Implementing DivSeek on a day-to-day basis (JFU)

The JFU operationalizes work plans and facilitates their implementation on a day-to-day basis. It develops drafts work plans, draft budgets, resource-mobilization plans and annual reports for consideration by the SC and the PA. It supports the definition of operational guidelines for implementing DivSeek; mobilizes resources for DivSeek work plans; and administers the JFU budget. It also organizes PA and SC meetings; deals with DivSeek membership requests and recruitment; promotes linkages to other initiatives of relevance for DivSeek; and assists the SC in collecting information about interactions among partners.

[16] DivSeek Charter [https://www.dropbox.com/s/aj9qcsa89pvyveh/DivSeek%20Charter.pdf]

Four organizations with global reach and complementary constituencies contribute one representative each to the JFU to jointly implement DivSeek: the *Global Crop Diversity Trust* (Crop Trust), the *Global Plant Council* (GPC), the *Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture* (International Treaty), and the *CGIAR Consortium Office* (CGIAR-CO) (Figure 6).

Each of these four organizations contributes a wide range of expertise and technical knowledge to the DivSeek JFU, and examples of some of these are provided in the diagram below.

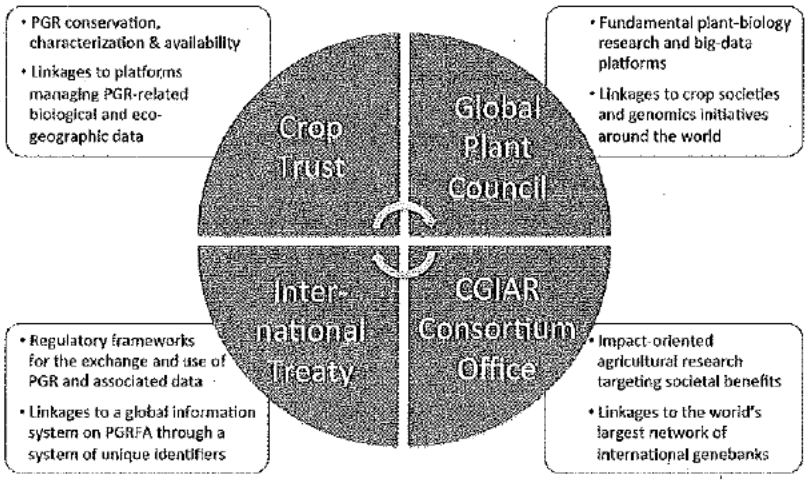


Figure 3: A non-exhaustive set of examples of how the Crop Trust, GPC, International Treaty and CGIAR-CO contribute to the Joint Facilitation Unit

Towards a DivSeek Program of Work

One of the first tasks of the Steering Committee will be the development of a program of work. To assist the SC, the JFU provide appropriate materials for elaborating a Program of Work, for example the JFU will prepare a 'menu' of suggested areas of DivSeek activities. Initially this will be based on the priorities identified by partners during the 2015 Assembly (outlined below), the priorities will compared to the goals of existing projects of relevance for DivSeek, which will be identified in a preliminary landscaping study (briefly outlined below) in advance of the Steering Committee meeting in May. This approach should allow us to suggest a portfolio of pilot projects that could provide 'building blocks' for DivSeek.

A second component of an initial Program of Work will deal with DivSeek management and other 'housekeeping' issues. This component builds upon the DivSeek governance framework. It will likely include activities such as:

- Convening Partner Assembly and Steering Committee meetings
- A communication strategy
- Building linkages with other relevant initiatives
- Refining operational guidelines for DivSeek governance bodies, for example, based on insights provided by the ASU research project

Priority Areas Identified by Partners

The survey conducted as part of the 2015 Partners' Assembly[17], as well as recent insights from discussions about implementing DivSeek in selected crop communities, point to the following priority areas of work:

- Information management
  - Data standards
    - Permanent Unique Identifiers (PUID) for germplasm and datasets
    - (Meta)data standards for high-density molecular data (GbS, SNP arrays)
    - (Meta)data standards for various types of phenotypic data
  - Software platforms
    - Data repositories and tools for distributing, processing and visualizing validated datasets
    - High-performance computation platforms and repositories for storing and processing primary data
    - Data bases and tools for gathering and managing primary data
- Community building & networking
  - Inventory of research approaches and lessons learned
  - Community-building: social network of PGR researchers
  - DivSeek as a neutral platform to broker germplasm and data exchange among organizations working on a particular crop
  - Joint advocacy efforts: public awareness and fundraising
- Research approaches & platforms
  - Accession-sampling strategies (inbred, outcrossing crops, etc.)
  - Genotyping platforms
  - Phenotyping methods
  - Targeting and quantifying intended impacts (crop improvement programs, upstream research, genebank operations, etc.)
- Rights management
  - Measures to incentivize sharing of germplasm for DivSeek research for PGR regulated under the International Treaty and the Convention on Biological Diversity
  - A commonly accepted open-data rights framework for sharing DivSeek data within and across projects
  - The role of public-private partnerships in DivSeek
- Capacity building
  - Training in data management, curation and analysis
  - Identification of other priority areas and targeted audiences for capacity building efforts

Landscape of Ongoing Projects

[17] Report on the technical session of the 2015 DivSeek Assembly https://www.dropbox.com/s/8eebwobfuiv6w6j/DivSeek%20Assembly%20report%20-%20technical%20session.pdf?dl=0

The JFU will conduct a preliminary landscaping study before the SC meeting in May. This study will take into account and expand upon previous efforts to identify 'DivSeek'-like projects across a variety of crops. We will attempt to identify a 'core collection' of ongoing projects by maximizing the 'diversity' according to three sets of criteria that can be used to describe individual projects:

- The importance of a particular crop for agriculture
- Factors that shape research approaches
- The regulatory framework(s) applicable to a crop and its wild relatives

Each of these three categories contains subcategories designed to broadly capture features of relevance to categorize different crops for the purposes of DivSeek (Figure 4).

The second area the landscape study will focus on the information-management domain. Data standards, data repositories and high-performance computational environments for data processing may be the three principal areas to investigate.

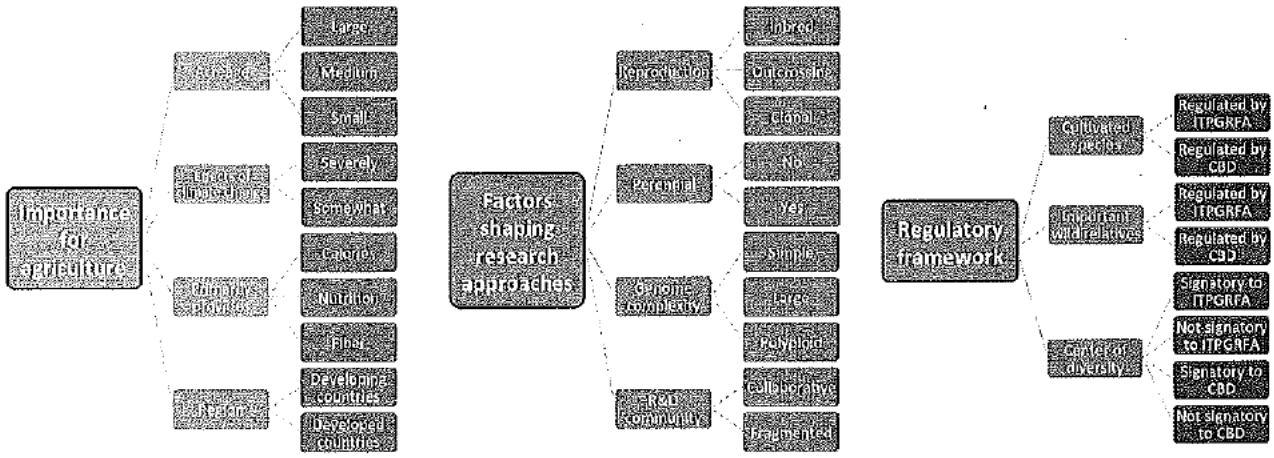


Figure 4: Possible set of criteria for choosing a subset of DivSeek-like projects across a variety of crops.

Bretting, Peter

From: [redacted] on behalf of Susan McCouch <srm4@cornell.edu>
Sent: Monday, May 04, 2015 11:58 AM
To: Andreas Graner (IPK); David Marshall; Elizabeth Arnaud (Bioversity); Emily Marden (UBC); Bretting, Peter; Rajeev Varshney (ICRISAT & GCP); Ruaraidh Sackville Hamilton (IRRI); Sarah Ayling (TGAC); Susan McCouch
Cc: Peter Wenzl; Daniele Manzella; Ruth Bastow; Wayne Powell (CGIAR CO); Susan McCouch
Subject: Introduction to DivSeek document

Dear DivSeek Steering Committee members,

Several of you have asked about the status of the 'Introduction to DivSeek' document discussed during our phone meeting on 28 April, and whether you are free to share it with interested parties.

At this time, that document should be considered a draft. We are interested in your comments and feedback so we can make changes accordingly. Until the document is finalized, we ask that you please do not circulate it.

We look forward to hearing from you about any changes you feel would improve it. The most helpful way to do this would be tp send the document around to SC and JFU members (cc'd above), with tracked changes, or tracked comments indicating specific edits, additions or requests for further clarification. Once we receive your comments, I will work with the JFU to integrate them and we will then bring the document back around for discussion during the in-person SC meeting on May 28 in Rome.

Once the document is finalized, our plan is to make it public over the DivSeek website.

In general, when DivSeek documents are first brought up for discussion in the SC, please treat them as drafts. Once these documents have been vetted through the SC and JFU, and your input has been integrated, my recommendation would be to make them public via the DivSeek website.

Using the DivSeek website to release DivSeek documents to the public will ensure that we have only one version of a document circulating, and that further comments and discussion surrounding that document can be dealt with through information channels established based on the governance structure and principles of DivSeek.

Please let me know if you agree with this recommendation. We are currently working to update and restructure the DivSeek website to improve navigability and make it easier for users to find information.

The SC will have a chance to review the updated web-site at our May 28 meeting in Rome, and your input will be essential as we endeavor to share information with our stakeholders and other interested parties.

Best regards,
Susan

--
Susan McCouch
Professor, Plant Breeding & Genetics
Cornell University
162 Emerson Hall

Ithaca, NY 14853-1901
Phone: +1 607-255-0420
Fax: +1 607-255-6683
Email: srm4@cornell.edu or mccouch@cornell.edu
Alternate Email: [redacted]

Bretting, Peter

From:	Peter Wenzl <peter.wenzl@divseek.org>
Sent:	Thursday, May 07, 2015 5:07 AM
To:	Andreas Graner; David Marshall; Elizabeth Arnaud (Bioversity-France); Emily Marden; Bretting, Peter; Rajeev Varshney (ICRISAT-IN); Ruaraidh Sackville Hamilton (IRRI); Sarah Ayling; Susan McCouch
Cc:	Daniele Manzella; Peter Wenzl; Ruth Bastow; Wayne Powell (CGIAR Consortium)
Subject:	Your bios
Attachments:	Bios of SC members.docx

Dear Steering Committee members,

As previously mentioned by Susan, we're in the process of updating the DivSeek website to reflect recent developments such as the election and upcoming meeting of the Steering Committee.

We're planning to add a page with a list of Steering Committee members and -provided you agree- would like to include links to your bios and photos. Our communications team has slightly edited your bios to bring them to a more similar format (see attachment).

Please let us know by next Monday (11 May), in case you prefer your bio or photo not be put online, or if you'd like to modify your bio.

Many thanks,

Peter,

on behalf of the Joint Facilitation Unit

--
Peter Wenzl
DivSeek Liaison
Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]
www.croptrust.org

Securing our Food, Forever

Elizabeth Arnaud

Bioversity International

Norwich Parc Scientifique Agropolis II,1990 Boulevard de la Lironde, 34397 Montpellier, France.

Elizabeth holds a MSc in Biology and a Master in Scientific data management techniques and communication. When she joined Bioversity, ELizabeth coordinated the Musa Germplasm Information System (MGIS) and then the CGIAR System-wide Information System on Genetic Resources (SINGER), and the development of the Bioversity geospatial database for collected crop samples. Since 2008, she has led the Crop Ontology project of the Generation Challenge Program, developed with CGIAR Research Centers and partners.

The Crop Ontology project is contributing to the Integrated Breeding Platform (IBP) and a new NSF-funded project called Planteome, led by Oregon State University, which aims to develop a set of reference ontologies for plants. Elizabeth is co-Principal Investigator of the Planteome project. She was the chair of the 2009 conference on Biodiversity Informatics Standards (TDWG) in Montpellier, France. Elizabeth is the head of Bioversity's delegation in the Global Biodiversity Information Facility (GBIF) Governing Body. Since 2012, she is also a member of the GBIF Scientific Committee.

Sarah Ayling

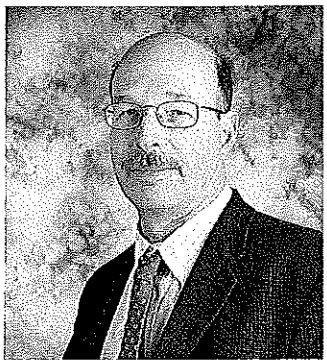
The Genome Analysis Centre (TGAC), UK

Norwich Research Park, Norwich, NR4 7UH, UK

Sarah leads the Crop Genomics and Diversity group at The Genome Analysis Centre (TGAC), which supports crop improvement through the application of genomics approaches. The group works on genome assembly, annotation, identification of variants and development of genomics tools and resources. Current species of interest include wheat, barley, rice, brassicas, pea, yam and forages.

Prior to joining TGAC in 2011, Sarah spent two years at the Centro Internacional de Agricultura Tropical (CIAT), Colombia, where she worked as a bioinformatician supporting the bean, cassava, rice and tropical forage programs. She also worked on the development of a laboratory information management system (LIMS). From 2008–2009 Sarah worked on the domestication of einkorn wheat at the University of Manchester, UK. And from 2005–2007, as a software developer within the Ensembl project, she managed and developed the automated genome annotation pipelines. Sarah completed her bioinformatics PhD in 2006 studying phylogenetic network approaches and visualization.

Sarah has a keen interest in making genebank materials more accessible to the community and produced a technical report on the feasibility of sequencing collections for the Global Crop Diversity Trust in 2012.



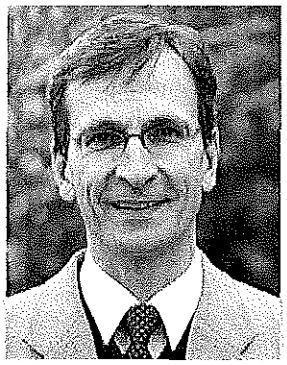
Peter Bretting

US Department of Agriculture–Agricultural Research Service (USDA–ARS), USA

Mailstop 5139, 5601 Sunnyside Avenue, Beltsville, MD 20705-5139, United States of America

Since 1998, Peter Bretting has been the USDA–ARS Senior National Program Leader for Plant Germplasm and Genomes. He provides expertise in the form of co-leadership, coordination, and direction of USDA–ARS's national program of crop genetic resources (US National Plant Germplasm System), genomics, genetics, bioinformatics, and breeding research conducted at more than 50 locations nationally, with an annual budget of approximately $165 million. Concurrently, he also provides expertise as co-Principal Investigator for the Global Crop Diversity Trust, Bioversity International, and a USDA–ARS joint project to develop GRIN-Global as an international standard information management tool for plant genebanks.

Prior to that, Peter was Research Leader and Coordinator at the USDA–ARS's genebank at Ames, IA, US, and Collaborator-Associate Professor of Agronomy and Botany, Iowa State University, US, where he led one of the largest genebanks of the USDA–ARS National Plant Germplasm System (50,000+ accessions), conducted research on crop genetics and genetic-resource management, and taught a graduate-level course on 'Plant Genetic Resource Management'.



Andreas Graner

Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany

Correnssstr. 3, D-06466 Seeland OT Gatersleben, Germany

Andreas Graner has a PhD in Plant Genetics from the Technical University of Münich, Germany. He has been the Managing Director of IPK since 2007. Prior to that, he chaired the International Barley Sequencing consortium, and was a Scientific Advisory Board Member for the Biodiversity and Genetic Resources unit at the Federal Ministry for Food and Agriculture (2006–2008). Among his other professional experiences, Andreas was Head of the Federal *ex situ* Genebank at IPK Gatersleben, Germany (1999), and Professor for Plant Genetic Resources at the University of Halle, Germany (1999).

As a molecular geneticist, Andreas has been involved in the development of genomics resources and deployment of molecular tools for crop genetics and breeding for over 30 years. Currently, he is actively involved in the establishment of Biodiversity Informatics at the IPK, which involves the convergence of classical genebank documentation with Bioinformatics and includes the development of software tools for the analysis and visualization of PGR-related data.

Andreas is convinced DivSeek will help build an infrastructure to systematically capture and store genomics and phenomics data, to be converted into high-value information that will enable informed access to PGR for research and breeding.

Emily Marden

University of British Columbia, Canada

School of Law, Allard Hall, 1822 East Mall, Vancouver, British Columbia V6T 1Z1, Canada

Emily Marden is a practicing attorney and Research Associate in the School of Law at the University of British Columbia (UBC), Canada. Her expertise is in governance related to genomics and agriculture. Emily has extensive experience addressing issues of intellectual property (IP), innovation policy and regulatory strategy in the biotechnology and agricultural arenas for public sector, government and the private sector stakeholders. In recent years, she has been working with the Rieseberg laboratory at UBC to address the impact of regulatory, IP and sharing mechanisms on research and innovation. Emily completed her undergraduate degree in molecular biology has graduate degrees in the History and Philosophy of Science, and Law.

David Marshall

James Hutton Institute, UK

Invergowrie, Dundee, DD2 5DA, UK

David Marshall has recently stepped down from leadership of the Information and Computational Sciences Group at the James Hutton Institute, UK. With a first degree in Botany and a PhD in Plant Population Genetics, his research career of over 30 years has focused on the development and deployment of molecular markers in plant genetics and breeding applications. His involvement with the international plant genetic resources community began during his tenure in the Genetics Department at the University of Birmingham, UK. His current research interests are built on the new opportunities for germplasm characterization that have arisen from advances in sequencing and genotyping technologies. His research group has an international reputation for the development of software tools that have enabled the facile exploration of complex molecular datasets. He plays a significant role in a number of international advisory boards and expert groups.

Ruaraidh Sackville Hamilton,

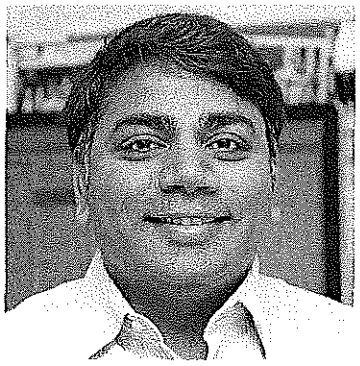
International Rice Research Institute (IRRI),

Philippines

Los Baños, Philippines

Ruaraidh Sackville Hamilton graduated in 1975 with a BA in Natural Sciences from the University of Cambridge, UK, majoring in plant genetic resources, plant breeding and plant ecology. In 1980 he was awarded a PhD, also from the University of Cambridge, on genetic diversity in a native forage legume species. Between 1980 and 2002 he worked mainly in Colombia, Wales and Italy on many aspects of plant genetic resources including plant breeding, genetics, ecology, statistics and database design. In 2002 he was appointed Head of the TT Chang Genetic Resources Centre (GRC) at the International Rice Research Institute in the Philippines, in which capacity he is responsible for managing the international rice genebank and the associated research on genetic diversity and bioinformatics, including the International Rice Informatics Consortium, one of the 'teeth' of the DivSeek 'comb'.

At IRRI, Ruaraidh has also taken a leading role in developing institutional policies and protocols that ensure not only compliance with the Treaty and but also a complete and effective value chain from genebank through breeder and farmer to consumer.



Rajeev Varshney

International Crop Research Institute for Semi Arid Topics (ICRISAT), India

Patancheru, 502 324, India

Rajeev Varshney is Research Program Director for Grain Legumes, and Director of the Center of Excellence in Genomics (CEG) at ICRISAT. He is also Winthrop Research Professor at the University of Western Australia. Previously, he served at the CGIAR Generation Challenge Program based in Mexico as Theme Leader for six years; and worked at the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany, for five years.

Rajeev has a background in molecular genetics and possesses about 20 years of research experience in international agriculture. His primary contributions include genome sequencing of pigeonpea, chickpea, peanut, pearl millet, sesame, mung bean and adzuki bean.

Rajeev has been a pioneer in the field of sequencing and re-sequencing of crop genomes and has been advocating the use of next-generation sequencing technologies to understand the genetic architecture of germplasm.

Rajeev has a prolific publication record with an *h*-index of 51 and more than 250 publications in leading journals of international reputes including *Nature* and *Nature Biotechnology,* among others; ten edited books; and special issues (as guest editor) for several journals to his credit.

Bretting, Peter

From:	Peter Wenzl <peter.wenzl@croptrust.org>
Sent:	Friday, May 01, 2015 3:48 AM
To:	Andreas Graner; David Marshall; Elizabeth Arnaud (Bioversity-France); Emily Marden; Bretting, Peter; Rajeev Varshney (ICRISAT-IN); Ruaraidh Sackville Hamilton (IRRI); Sarah Ayling; Susan McCouch
Cc:	Daniele Manzella; Peter Wenzl; Ruth Bastow; Wayne Powell (CGIAR Consortium)
Subject:	Minutes for DivSeek SC conference call on April 27th
Attachments:	Minutes of SC conference call (27 April 2015).pdf

Dear Steering Committee members,

Please find attached the minutes for our conference call earlier this week.

For those who couldn't attend or would like to re-listen, you can download an audio recording of the call until May 5th
here: [redacted]

Please don't hesitate to contact us in case you've any questions or would like us to clarify any particular aspects.

All the best,

Peter, on behalf of the Joint Facilitation Unit

--

Peter Wenzl
DivSeek Liaison
Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]
www.croptrust.org

Securing our Food, Forever

DivSeek Steering Committee conference call

27 Apr 2015

Present

- Steering Committee
  - Susan McCouch (Chair)
  - Elizabeth Arnaud
  - Sarah Ayling
  - Andreas Graner
  - Emily Marden
  - David Marshall
  - Ruaraidh Sackville Hamilton
- Joint Facilitation Unit
  - Ruth Bastow
  - Daniele Manzella
  - Peter Wenzl

Apologies

- Steering Committee
  - Peter Bretting
  - Rajeev Varshney
- Joint Facilitation Unit
  - Wayne Powell

Meeting minutes

- ***Agenda item: Introduction and welcome***
  - *Susan* welcomes Steering Committee (SC) members and notes their diverse range of expertise. She reminds the SC that DivSeek is still in its early phases and will need continuous support and guidance from the SC members to be successful. This initial gathering of the SC is to help define who we are and where we are going.

- ***<u>Agenda item</u>: Developments since the first Assembly: updates from Joint Facilitation Unit members***
    - *Daniele*: International Treaty Secretariat
        - Circulated among attendees of the January Assembly in San Diego a meeting report, together with the adopted DivSeek Charter and list of organization that attended.
        - Signatories of last year's expression of interest who did not attend the Assembly were also emailed the Charter and asked to endorse it. Responses have been limited, and follow-up on a case-by-case basis is required.
        - Before the Assembly in San Diego, the first Consultation on the Global Information System on Plant Genetic resources for Food and Agriculture (COGIS), in accordance with Article 17 of the International Treaty was held. Amongst others, the topic of permanent unique identifiers (PUID) was discussed. A subsequent task force concluded that Digital Object Identifiers (DOI) was the option preferred by most.
        - One of the multi-country funded projects of the 3rd project cycle of the Benefit Sharing Fund, led by the Indonesian Agency of Agriculture Research and Development (IAARD) in close collaboration with IRRI, will use PUID and develop an API to link rice data sets within the Multilateral System.
    - *Peter*: on behalf of the CGIAR CO
        - A paper outlining future options for funding CGIAR genebank operations after the conclusion of the current genebank CRP, has been drafted by the CGIAR Consortium Office, the Crop Trust, and genebank managers from centers. The paper concludes with a chapter covering future developments that references DivSeek.
    - *Peter*: Crop Trust
        - The Crop Trust-coordinated Crop Wild Relative Project will fund the implementation of the Germinate data repository from the James Hutton Institute (JHI) for rice and sunflower data. The platform has previously been used for barley and potato data at JHI. It has also been implemented in CIMMYT's Seeds of Discovery project for maize and wheat data.
        - Peter attended a meeting by a small group representing the global soybean genetic-resource community, organized in Seattle by Scott Jackson. The objective was to develop a joint action plan for a DivSeek-like project in the crop. DivSeek was seen as an option for could act as neutral broker for issues surrounding data and germplasm exchange.
        - Roberto Papa from from Università Politecnica delle Marche in Ancona (Italy) contacted DivSeek regarding a recent ERA-CAPS, NSF funded project that will genotype 10,000 globally sourced

accessions of common bean. Roberto requested to discuss how to associate this project with DivSeek.

- *Ruth*: GPC
  - Ruth updated the SC on a topic not covered in the introduction paper provided for this meeting: an ERA-Net on a Coordinating Action in Plant Sciences (ERA-CAPS) is bringing together funders from across Europe, Canada, New Zealand and the US (NSF). ERA-CAPS has formed a taskforce on data standards in plant molecular sciences, which will produce a report of relevance for DivSeek in coming month(s).
- Additional comments and discussions
  - *David*: a number of relevant projects funded by the Bill and Melinda Gates Foundation (BMGF) are working towards commonalities in APIs for breeding related data. There will be common ground between these efforts and DivSeek, e.g., standards for passing information such as molecular markers, trait info and line data.
  - *Susan*: there is a global working group on APIs, which DivSeek might consider interacting with; the SC may wish to discuss this further.
  - *Peter*: recently met with representatives of BMGF and discussed a potential involvement of the Foundation in DivSeek, given their ongoing re-assessment of funding priorities. The BMGF-funded GOBII project deals with high-density molecular data and thus is very relevant to DivSeek. Given previous investments in agricultural data management (IBP, GOBII), further investments in adapting GOBII tools towards a genebank context, are unlikely unless matching funding can be identified from elsewhere first.

- ***Agenda item: Governance-related issues***
  - *Susan* reminds the SC that the focus of DivSeek is to promote the use of genetic diversity for food and nutritional security and other economical benefits. The SC needs to consider how DivSeek via supporting, cross inking and adding value to current activities can accelerate crop improvement.
  - The key principles of DivSeek as clearly outlined in the DivSeek introduction paper distributed in advance to the meeting. DivSeek requires a form of internal governance that enables it to follow these principles and function effectively.
  - There are three components of the current governance structure:
    - The Partners' Assembly (PA) is the ultimate decision body of DivSeek;
    - The Steering Committee (SC), which guides the implementation of the DivSeek strategy;

        - The Joint Facilitation Unit (JFU) serves both the PA and the SC.
    - These 3 different structures need to work in synchrony with each other to ensure DivSeek succeeds. The role of the SC is to advise and guide the implementation of the strategy outlined by the PA, and to work closely with the JFU to achieve this.
    - Comments & discussion:
        - *Emily*: unclear ref. (a) the role of the JFU in setting the agenda, (b) where the initiatives come from and (c) how this actually all works from day to day. What is the balance between a member-driven organization and the JFU that deals with the day-to-day issues? Have these roles been determined or is this still something that we are planning to discuss and determine?
        - *Susan*: the SC is only just getting under way, and we can now work towards achieving a balance among the three components of the governance structure.
        - *Daniele*: at the PA, concerns were raised over the role of the JFU. The Charter helps to clarify some of these. The JFU facilitates implementation of a program of work and responds to the instruction from the PA. In between PAs, the JFU responds to the instructions of the SC. The JFU is currently drawing up an operational plan that will be submitted to the SC at the May meeting. This document will clarify how the JFU organizations interact and function together (e.g. lines of communications, responsibilities, etc.). The operational plan will be submitted to the 2016 PA.
        - **Action point:** the JFU to submit a draft operational plan to the SC in advance of the meeting in May. This plan will describe the roles of the four JFU organizations and outline how the JFU can operate in an efficient and effective manner to respond to a rapidly changing external environment.
        - *Elizabeth*: What are the stages/timeline in the development of the program of work (work plan)?
        - *Daniele*: based on the feedback collected from Partner organizations, the "DivSeek – an introduction" document presented in this SC call and the advice received from the SC, the JFU will develop a first draft programme of work for the SC to discuss in May; a revised draft will be submitted to the PA for consideration and approval
        - *Andreas*: What is the intersection between the work plan and financial support for the work plan? Is the first step to define a work plan, get this agreed by the PA and then use this to attract funding or vice versa?
        - *Peter*: The JFU will provide ideas on what the work plan could contain based on the feedback from the PA survey; but it is up to

the SC to determine the contents of the work plan that will be submitted to the PA in 2016. The work plan could be presented to the PA with some suggestions about costs and funding priorities.

- *Susan*: A number of projects that have been discussed to date are already funded. At present DivSeek doesn't have funds to support projects but instead can consider projects that already exist and assess where coordination and interaction across projects (via DivSeek) would be beneficial to leverage information and add value.
- *Andreas*: A number of organizations/foundations are interested in funding DivSeek e.g. Syngenta Foundation, Pioneer and BLE. How will this be dealt with? Is there a framework in place to allow this happen? How would DivSeek manage such opportunities?
- *Peter*: DivSeek perhaps could consider the development of a "landing path" for those that wish to align themselves and link up with DivSeek, to be able to receive funds that are targeted toward the strategic goals of DivSeek.
- *Susan*: DivSeek will need to manage expectations and relationships with the private sector. The SC needs to define how private sector organizations can join as partners and what is the role of the private sector and private sector funds.
- Action point: Interaction with, and role of the private sector to be discussed at SC meeting in May.*Daniele*: The JFU is communicating with private sector partners that attended the PA. The Charter foresees the development of operational guidelines on private sector engagement. The ASU project on institutional and organizational factors for enabling data access exchange and use aims for DivSeek (which the Treaty Secretariat ad the Crop Trust have agreed to co-fund) will also provide recommendations regarding private sector engagement, based on an analysis of comparable experiences within and outside the plant sector. Regarding DivSeek's program of work, it would be useful for the SC to discuss whether to share the program of work (or sections thereof) with industry for consultation.
- *Dave*: There's a spectrum between public and private organizations (e.g. Palm Oil Board in Malaysia) → we need to lay out a framework that caters for a multiplicity of partners along the public to private spectrum
- Action point: Can SC members please provide examples of consortia or similar efforts that have dealt with this public-private continuum?

- ***Agenda item: Looking forward: landscaping study***
  - *Ruth*: Plan is to provide a small, preliminary landscaping study for May

meeting → identify potential 'building blocks' for DivSeek
    - 3-D space spanned by three types of 'selection' criteria: importance of a crop for agriculture, factors shaping research approaches, regulatory framework (see Figure 4 in the "DivSeek – an introduction" document)
    - Similar approach to define criteria for projects in the data-management domain
    - Action point: Could SC members provide feedback ref. suggested 'selection' criteria and projects that would cover a diversity of factors displayed?
    - Move forward by prioritizing and focusing on efforts that DivSeek (as a currently unfunded initiative) should build upon to offer value to its members

- ***Agenda item: Looking forward: draft program of work***
    - *Peter:* Suggested approach is to identify relevant ongoing projects through the landscaping study that Ruth has outlined, and intersect them with priority areas identified by DivSeek partners through the survey distributed during the January Assembly
    - Priority areas outlined:
        - Data/information management, including data standards and software platforms
        - Community-building and networking (e.g. bringing together crop communities to stimulate data/germplasm sharing)
        - Research approaches and technology platforms used by different groups for different crops
        - Rights management (data sharing, public-private partnerships)
        - Capacity-building: not prioritized in survey, but perhaps worthwhile considering upfront, in the light of the germplasm vs. technology providers schism?

- ***Agenda item: Membership: draft letter of interest***
    - *Daniele*: Several organizations have contacted us via the DivSeek website
    - JFU has drafted a form that includes a number of questions for prospective partner organizations to answer → requests then to be forwarded to SC in order to select relevant organizations
    - Comments:
        - *Dave*: Form should include a brief description of what DivSeek is
        - *Ruaraidh*: Use "desired" or "anticipated" instead of "intended" when talking about contributions by prospective DivSeek partners
        - *Andreas*: Ask prospective partners to state their commitment to actively contribute to DivSeek
        - *Daniele*: this commitment would actually be formalized by endorsing the Charter

- Action point: Form to be finalized, based on comments that the SC members are asked to send in writing before the meeting on 28 May.

- ***Agenda item: Preparation for Steering Committee meeting on 28 May***
  - Invitation for dinner on 27 May; travel and hosting arrangements are in course of completion
  - Action point: JFU to provide a draft meeting agenda in advance to the meeting
  - Action point: JFU to provide the following background documents in advance to the meeting, including:
    - A preliminary landscaping study
    - A document with elements for a draft program of work
    - A documents with elements for a draft resource mobilization plan to accompany the draft program of work, and
    - A document outlining an operational plan for the JFU

Bretting, Peter

From: Peter Wenzl <peter.wenzl@croptrust.org>
Sent: Thursday, April 30, 2015 7:15 PM
To: 이석하
Cc: Scott Allen Jackson; Jones Jennifer; Abberton Michael (IITA); Cannon Steven [AGRON]; Joost Richard; Nelson, Randall - ARS; Okamuro, Jack; 邱丽娟 QIU Lijuan; 李英慧; Masao Ishimoto; Akito Kaga; Bob Stupar; Ricardo Abdelnoor; Bretting, Peter; marcelofernandes.oliveira@embr; Cannon, Steven; François Belzile
Subject: Re: Follow Up: Glycine meeting Seattle, April 15-17

Dear all,

Time's running and it's already two weeks since the meeting in Seattle! I wanted to thank you -and Scott in particular- for having invited me to your meeting. I really enjoyed it and there was a lot I leaned from your presentations and the ensuing discussions.

If you don't mind, I would like to report back to the DivSeek Steering Committee some of the key take-away points from the meeting, such as:

- Your plan to assemble a cross-genebank accession inventory
- Your approach to attempt to integrate existing data sets through a community-wide catalogue of GBS sequence tags identified across a broad range of accessions
- The plan to genotype all existing core collections on a single platform to create a global picture of the total diversity available, which could then be used to place individual accessions into a global context
- Your plans to explore opportunities of working with DivSeek in the area of software tools, databases, and hosting large datasets.
- Your quest to at least achieve data-sharing for cases where germplasm sharing is difficult.

In case there're any other important items I've missed here, which you would like me to report back to the Steering Committee, please let me know.

Looking forward to a continuing relationship,

All the best, Peter

--
Peter Wenzl
DivSeek Liaison
Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]
www.croptrust.org

Securing our Food, Forever

duplicate email trail removed

Bretting, Peter

From:	Peter Wenzl <peter.wenzl@divseek.org>
Sent:	Tuesday, February 24, 2015 8:00 AM
To:	info@divseek.org
Cc:	Susan McCouch; Ruth Bastow; Daniele Manzella; Powell, Wayne (CGIAR Consortium)
Subject:	Report on the technical part of DivSeek meeting in San Diego
Attachments:	DivSeek technical meeting - Annexes.pdf; DivSeek technical meeting - Summary.pdf

Dear DivSeek colleagues,

Please find attached a report on the technical part of the DivSeek meeting in San Diego this January.

The first part of the report includes a brief summary and a side-by-side comparison of the projects and platforms presented by the speakers, to identify potential commonalities and complementarities among ongoing efforts.

The presentations themselves can be downloaded as a zip file here: [redacted]

The second part of the report analyses the detailed written feedback we received from you via the questionnaire that was distributed during the meeting.

Your guidance in the form of previous surveys and the feedback to this questionnaire will provide a basis upon which the incoming *Steering Committee*, with support from the *Joint Facilitation Unit*, will develop an initial program of work for DivSeek.

You will receive an invitation to elect *Steering Committee* members within the next several days.

Best wishes,

Peter
(on behalf of the *Joint Facilitation Unit*)

--
Peter Wenzl
DivSeek Liaison
Global Crop Diversity Trust
Platz der Vereinten Nationen 7
53113 Bonn, Germany
Office: +49 228 85427 126
Mobile: [redacted]
www.croptrust.org

Securing our Food, Forever

Annex A:

Summary of presentations given at the DivSeek meeting in San Diego (9 Jan 2015)

Table 1: *Crop-focused germplasm-characterization efforts*

	Kevin Pixley (CIMMYT)	Ruaraidh Sackville Hamilton (IRRI)	Nils Stein (IPK)	Rajeev Varshney (ICRISAT)	Theo van Hintum (CGN)	Wayne Powell (CGIAR CO)
Name of project or program	Seeds of Discovery (SeeD)	International Rice Improvement Consortium (IRIC)	Biodiversity Informatics to Close the Gap from Gene Information to Educated Utilization of Diversity Stored in Genebanks (BRIDGE)	Multiple projects for different crops	National genebank example	CGIAR Consortium Office
Crop(s)	Maize, wheat	Rice	Barley	Sorghum, chickpea, pigeon pea, groundnut, pearl millet	-	Rice, banana, plantain, beans, cassava, tropical forages, maize, wheat, potato, sweet potato, Andean roots & tubers, grain legumes, barley, forage & range crops, trees, dryland cereals, grain cereals, cowpea, cassava, yam
Type of effort	Research project	Consortium	Research project	Several research projects	Genebank operations	Consortium of research centers
Objectives	Identify & mobilize useful diversity in maize & wheat genebanks	Provide information and tools to accelerate rice breeding by exploring & harnessing rice diversity	Flagship project for DivSeek & blueprint for IPK genebank to facilitate informed utilization of genetic resources and improve conservation management; follows outline of SeeD project	To develop genomic resources for ICRISAT's "orphan" crops and characterize their genetic resources	-	-
Main areas of activity	[1] "Core" activities (genotyping, databases, IP strategy); [2] "associated" activities building upon core activities (specific traits and geographic areas)	[1] Organize genotypic, phenotypic, expression and other data into information system for global rice research community; [2] Provide user-friendly access to data through a single portal	[1] Catalogue diversity; [2] upload new and legacy data, in standardized formats, to data warehouses for dissemination via intuitive user interfaces; [3] address population-genetics and domestication-related scientific questions	[1] Develop genomic resources; [2] characterize genetic diversity; [3] phenotype traits targeted by breeders; [4] perform GWAS	-	-

	Kevin Pixley (CIMMYT)	Ruaraidh Sackville Hamilton (IRRI)	Nils Stein (IPK)	Rajeev Varshney (ICRISAT)	Theo van Hintum (CGN)	Wayne Powell (CGIAR CO)
Progress & highlights	Genetic analysis of almost entire maize and 1/3 of the wheat collections at CIMMYT completed; ICARDA's wheat collection to be genotyped in 2015; millions of phenotypic measurements; implementing Germinate as a data repository; developing 'bridging germplasm' half way between genebanks and breeding programs; capacity-building underway	4,500 available lines available through genebank; whole-genome re-sequencing data for approx. 3,000 purified lines; characterization data from genebank; not much other phenotypic data yet (no large-scale phenotypic evaluation data); some expression data	To initiate in May 2015	Example for chickpea: 300 accessions whole-genome resequenced -> identified 4.7M SNP; established field-based, high-throughput phenotyping systems including for root traits; identified marker-trait associations for yield, etc.; now whole-genome resequencing of 3,000 accessions	-	-
Challenges encountered	55 funding reduction in 2013; SMTA is obstacle for engaging with private partners; some partners hesitant to contribute their genebank materials to project	No funding for initiative; how to structure PPP; currently no issues with SMTA, but this may change in the context of hybrid rice; inadequate genomic references (data quality become a function of the genetic distance from single reference genome); resources for big-data platforms	Funding for phenotyping	Data storage; germplasm sharing	Managing large numbers of accessions; underfunded; genebanks lack knowledge of genomics and bioinformatics; one-way flow out of genebanks into research community without information flowing back and adding value to genebanks and most derived materials getting lost	-
Subsetting of accessions	Genetic characterization: all accessions (27,000 maize, 120,000 wheat); phenotypic characterization: subsets of varying size	Genetic characterization: subset of 3,000 accessions	Genetic characterization: all accessions (20,000)	Genetic characterization: Core collections (10) and min-core collections (0.4) derived from a total of 110,000 accessions of five species; reference sets of 300 accessions/crop defined based on SSR using the "composite" collection approach (GCP); starting with subsets and moving towards systematic genetic characterization	-	-
Sample preparation for genetic analyses	Pool of 30 seeds/accession (maize); single seed/accession (wheat)	Purified lines derived through SSD from genebank accessions	Still under debate; pool of 50 seeds per accession or single seed/accession	-	-	-

	Kevin Pixley (CIMMYT)	Ruaraidh Sackville Hamilton (IRRI)	Nils Stein (IPK)	Rajeev Varshney (ICRISAT)	Theo van Hintum (CGN)	Wayne Powell (CGIAR CO)
Type of genetic data generated	Genotyping by sequencing (DArTseq & Cornell platforms)	Whole-genome resequencing with 10 million SNP; SNP-chip data (700K and 44K)	Genotyping by sequencing; whole-genome shotgun sequencing on selected accessions	Genetic characterization: create reference genome sequences alone or in collaboration (pigeon pea, chickpea, groundnut, peanut); whole-genome resequencing of subsets (e.g., chickpea); genotyping by sequencing (Cornell platform)	-	-
Type of phenotypic data generated	Large-scale field trials to evaluate heat/drought tolerance, disease resistances, and grain-quality traits	Limited phenotypic data to this date; gene-expression data	Make existing legacy data available; possibly phenotype selected accessions	Field trials in African and Indian locations; drought tolerance (incl. root traits); application of field-based, high-throughput techniques	-	-
Population development?	"Bridging germplasm"	no	no	-	-	yes; boundaries between pre-breeding and breeding are increasingly blurred
Access to germplasm	SMTA for exchanging germplasm accessions	SMTA for exchanging germplasm accessions; question: are SSD-derived lines "germplasm under development"?	-	-	-	SMTA for exchanging accessions
Controlled vocabulary and ontologies used?	-	-	Plans to implement community standards	-	-	-
Type of information systems used	Germinate as data repository; working with, and testing databases of IBP and DArT for suitability for managing primary data	Oracle database as data repository; Galaxy workbench for analyzing genetic data; discussions with iPlant and National University of Singapore to identify other already existing resources	Link to IPK genebank information system; will create data warehouse by building on existing solution where possible	Currently in-house; now talking to iPlant	-	Various

	Kevin Pixley (CIMMYT)	Ruaraidh Sackville Hamilton (IRRI)	Nils Stein (IPK)	Rajeev Varshney (ICRISAT)	Theo van Hintum (CGN)	Wayne Powell (CGIAR CO)
Governance & data-sharing	CGIAR/CIMMYT open-data access policies; project partners get 2-year period of initial, preferential access to data for publishing; PPP to access know-how & tools	Consortium; contributing partners get early, preferential access to data and tools under development by others (= incentive); others get early access to contributing partners' data and tools (= disincentive to make everything immediately publicly available); most public partners to contribute phenotypic data; private partners have to contribute funds; compliant with CGIAR Intellectual Assets Principles as much as possible; all released data linked to purified lines (available through genebank) and will eventually be made publically available	Sharing of genomic data as soon as possible and without strings attached	Data (allele calls) available as global public goods		CGIAR Intellectual Assets Principles
How could DivSeek add value to project/program?	Forum for discussing & sharing experiences, approaches, methods; advocating best practices for data management; assure interoperability between projects; facilitate access to informatics tools; options to host big data sets; contribute to create pre-competitive domain	IRIC = DivSeek for rice (a "tooth in the DivSeek comb"); chance to secure funding for a "mega project" to mine genetic resources;	Platform to link with other projects to adapt already proven solutions particularly for information management	Data analysis, storage & hosting; assist with germplasm sharing issues	Ensure that genebank managers and genomics specialists work together to design studies and select materials; ensure that genomics data are linked back to accessions and add value to genebank collections to better serve all types of future users; convincing use cases and pilot projects; intuitive interfaces to genomics data; capacity building;	Genebanks are part of the CGIAR's core business and should be involved in DivSeek to get them to the next level

	Kevin Pixley (CIMMYT)	Ruaraidh Sackville Hamilton (IRRI)	Nils Stein (IPK)	Rajeev Varshney (ICRISAT)	Theo van Hintum (CGN)	Wayne Powell (CGIAR CO)
How could project/program contribute to DivSeek?	Sharing experience with inbred and outcrossing crop in terms of genotypic analyses, data storage; vision on how to ensure equity (through capacity building and other means)	Rice = "Arabidopsis with meaning": model for deploying genomic data in DivSeek projects: what's going to be possible for other crops, can be done now in rice -> learning to benefit other DivSeek projects	Can contribute to DivSeek through IPK's experience with genebank information management systems, bioinformatics and population genomics		Provision of genomics-ready materials (SSD and research populations derived from accessions; etc.)	The CGIAR genebanks are the backbone of the international genebank system and hence would provide a substantial portion of the accessions to be characterized; DivSeek could be part of the call for new CGIAR Research Programs (CRP)

Table 2: *Information management-related efforts*

	Laurel Cooper (OSU)	David Marshall (JHI)	Matthew Vaughn (TACC)	Susan McCouch (Univ. Cornell)	Paul Kersey (EMBL-EBI)
Name of project or program	Planteome: Common Reference Ontologies and Applications for Plant Biology (www.planteome.org)	Germinate	iPlant	Genomic and Open-source Breeding Informatics Initiative (GOBII)	ELIXIR
Crop(s)	Most data currently is from maize, Arabidopsis and rice, along with 20 other species; expanding towards all green plant species, particularly crops	Barley, maize, wheat, potato	Any	Rice, maize, wheat, sorghum, chickpea	-
Type of effort	International collaboration to develop centralized platform	Ongoing user-driven software-development effort	10-year software-development project	Software-development project	Provider of bioinformatics services and databases for life sciences
Objectives	Provide centralized platform for plant ontologies, annotated molecular datasets, and visualization & analysis tools for integration in online plant resources; co-develop metadata standards	Provide a multi-crop data repository/warehouse for genotypic, phenotypic, pedigree & climatic data to characterize germplasm	Design, develop and operate a modular & extensible cyber-infrastructure (services, hardware, software) for computational biology in all life sciences (beyond plants) to enable discovery, cloud computing and data storage	Put in place systems, databases, tools and analytical pipelines so plant breeders can routinely apply high density genomic information in cultivar development	Pan-European research infrastructure for biological information which maintains data beyond the duration of projects
Main areas of activity	[1] Consolidate, expand and maintain plant reference ontologies; [2] develop standards, workflows and tools for ontology development, curation and annotation; [3] provide a portal for ontology-annotated plant genomics data and germplasm; [4] develop semantic query, analysis, and visualization tools for curation and community-based annotation; [5] outreach and training in tools and ontologies	[1] Develop, expand & implement the data repository; [2] develop and link accessory data-query, visualization and analysis tools	[1] Enable safe & secure data storage and data-sharing; [2] provide access to high-performance applications; [3] provide access to common datasets	[1] Build database + API for high-density genotypic data; [2] link genomic database to pedigree database to interpret data in breeders' terms (IBD, haplotypes, etc.); [3] develop breeding schemes and pipelines to incorporate genomic selection in breeding programs; [4] build capacities	-

	Laurel Cooper (OSU)	David Marshall (JHI)	Matthew Vaughn (TACC)	Susan McCouch (Univ. Cornell)	Paul Kersey (EMBL-EBI)
Progress & highlights	Recently funded by NSF; grew out of the Plant Ontology and the Gramene project; Plant Ontology collaborates with more than 50 projects and 1,200 terms defined which are linked to 2.2M data points for 23 species (mainly maize, rice, Arabidopsis); Gramene collaborates world over with major plant bioinformatics resources	Currently version 3 (multi-language support; 20-fold quicker data extraction due to Java re-write); has authentication system, interface for smartphones; accessory software includes tools to visualize genetic diversity (3D PCA/PCO plots), graphical genotypes, pedigrees, and comparative maps; graphical genotypes integrated with phenotypic data and linked to 3D diversity visualizer	1/3rd of funding each dedicated to training, outreach and support; currently 1 petabyte of user data; access to services provided through API in a manner that supports the concept of federation; SoyKB: comprehensive genomics/molecular breeding web resource; genomic variation of 1,000 lines; partnerships with DOE, BBSRC, USDA	Just starting	Included the wheat IWGSC genome survey sequence in Ensembl (EMBL-EBI's genomics-data service); identified between-subgenome variants; implemented RESTful APIs for automatic data access
Challenges encountered	Integration and application of species-neutral reference vocabularies is challenging; our annotated datasets and education & training have stakeholder to recognize that ontologies enable comparison of data across studies or resources; the widespread adoption and integration in major national and international databases require semi-automated (trained curators) and automated (machine learning) curation workflows and extensive outreach and education	Data-quality issues not always fully recognized by users; missing GbS data; integration of data from multiple sources and platforms; scalability of data storage and analysis approaches; establishing equivalence of lines, loci, phenotypes, etc.	-	-	Polyploid genomes with sequences not contained in the reference; limited quality of known functional information; non-standard representation of individuals, accessions and phenotypes
Controlled vocabulary and ontologies used?	Yes: Plant Trait Ontology (TO), Plant Stress Ontology (PSO) and Plant Environment Ontology (EO), in addition to Plant Ontology (PO); will also include relevant aspects of Gene Ontology (GO), Cell type (CL), Chemical Entities (ChEBI), Protein Ontology (PRO) and the Phenotypic Qualities Ontology (PATO)	-	Community-dependent	-	reliance on controlled vocabularies

	Laurel Cooper (OSU)	David Marshall (JHI)	Matthew Vaughn (TACC)	Susan McCouch (Univ. Cornell)	Paul Kersey (EMBL-EBI)
Type of information systems used	Online portal with ontology browser, visualization tools and APIs; interacts with many external systems; backend: databases with ontology terms and annotations	Online portal with data Java-based warehouse in the background and associated analytical software tools	Online portal that provides access to a high-performance computing, data-storage and data-sharing environment	Backend DNA-sequence database linked, via API, to analytical tools that are pipelined to provide high-level functions such as genomic prediction, haplotype-tracking and decision support functions	DNA sequence databases
Governance & data-sharing	All data generated by project will be publicly available under a Creative Commons license based on attribution; data provided by collaborators may carry their licenses	Authentication system for different levels of access for different user types	Facilitate publication of data (links to DOIs, etc.); hands-off policy ref. IP, but encourage open-data policy; free data hosting in expectation that data will be made publicly available	-	All data freely available for any purpose; encouraging pre-publication submission of data and compliance with Toronto guidelines for early data release; data generated have time-limited monopoly for analyzing whole data sets; working with funders to normalize data sharing
How could DivSeek add value to project/program?	DivSeek would contribute to the development and adoption of Planteome resources; we are especially interested in the annotation of genetic variants, their phenotypic consequences, and source germplasm phenotype and passport metadata	-	Share DivSeek data and software tools with iPlant community; include DivSeek metadata standards iPlant Data Commons; engagement with genebank managers and breeders	-	-
How could project/program contribute to DivSeek?	Can contribute common standards and descriptors, starting with phase-1 studies; contribute to developing data analysis and mining tools and approaches; interested in exploring genetic-diversity domain, annotating phenomics data; training, education and outreach	An online data repository for dissemination of DivSeek data	Provide computational environment & support; distribution mechanism for large & complex data sets	Provide a backend, "under the hood" solution to storage of high-density genotypic data linked to a pedigree framework that facilitates tracking of IBD blocks	Provide access to DivSeek diversity data and phenotypic data linked to variants; maintain sequences of core reference sets (but not individual accessions); possibly: manage variant data for certain species & provide computational framework for comparative genomics; coordinate European efforts

Table 3: *Governance and funding*

	Eric Welch & Selim Louafi (ASU)	Steve Visscher (BBSRC)	Cindy Bell (Genome Canada)
Name of project or program	-	BBSRC	Genome Canada
Crop(s)	-	-	All crops are eligible for funding
Type of effort	Analysis of existing efforts in the DivSeek domain	Funding research	Funding research
Objectives	Identify preferred options for (i) governance and management of DivSeek, (ii) elements for a successful governance framework for DivSeek, and (iii) exchange, sharing and use mechanisms for pooled informational resources, and operational guidelines/principles	Fund research and post graduate training sin bioscience	Provide genomics solutions to sectors (including agri-food) that are vital to Canada's economic future
Main areas of activity	[1] Identify cases of existing projects; [2] analyze 3-4 cases for in-depth study through interviews, analysis of materials, collection of transactional access, exchange and contribution data, assessment of collaboration outcomes and analysis of survey results	[1] Basic research in food security, agriculture and bioscience for health (including nutrition), amongst other areas; [2] development of new 'tools' and national capability resources; [3] support International partnerships and cooperation	Support discovery and applied research as well as development of, and access to leading-edge technologies in the following areas: Agri-food, energy, environment, fisheries & aquaculture, forestry, health and mining
Progress & highlights	-	Wheat: contributed to global Wheat Initiative with (i) well-defined links among researchers, funders and the private sector, and (ii) expert working groups for key areas, including genetic resources and phenotyping; contributed to International Wheat Yield Partnership (IWYP)	Contribute to IWGSC through support of Curtis Pozniak's work; ongoing competition in agri-food sector which could provide funding for projects feeding into DivSeek
Challenges encountered	Behavioral impediments to data & material access, exchange and use	Too much work in silos in the past	-
How could DivSeek add value to project/program?	-	A platform to form strategic, international partnerships in food security, aligned with a general trend towards 'big science'; need for standards to enable interoperability to create synergies and optimize investments	-
How could project/program contribute to DivSeek?	Examine how other similar efforts have successfully pooled material and information for knowledge generation and innovation to inform DivSeek about how policies, rules, incentives, norms, preferences and technologies affect behavior; identify key design variables for DivSeek governance; identify ways to deal with competing objectives of partners through organizational design	Funding selected areas or topics	Provided support for initiating activities; support DivSeek's principles; participate in DivSeek's governance; funding of large-scale research projects that support DivSeek's mission

Annex B:

Categorization of feedback received during the DivSeek meeting in San Diego (9 Jan 2015)

Category	Fitting topics mentioned (% hits)[1]	Subcategories	How can DivSeek help you? (% hits)[1]	How can you help DivSeek? (% hits)[1]	Three top priorities for DivSeek? (% hits)[1]	Sum of three questions (% hits)[1]	Topics mentioned (in descending order of % hits)[1]
Information management	50.2	Data standards & procedures	4.4	1.3	12.2	17.9	Data standards & controlled vocabularies for raw data and analysis results including species-independent minimum standards (11.7); direct link of data to accessions and germplasm inventories (2.2); meta-data annotation for meta-analyses and date mining (1.3); working group to link data standards with other initiatives (0.9); data-quality annotation (0.9); data-acquisition procedures (0.9)
		Data bases & repositories	3.1	2.2	3.9	9.2	Tools for data storage and dissemination (7.4); go-to portal for crop diversity data (0.4); sustainability and longevity of repositories for analysis results (0.4); data repositories for small crops (0.4)
		Distributed data storage	0.9	1.7	3.9	6.5	Enable interoperability of distributed data bases/repositories through APIs & web services (6.1); interoperability with pre-breeding & germplasm enhancement efforts (0.4)
		Centralized data storage	0	0.4	1.7	2.1	Hosting of data in centralized storage platform (2.2)
		Data analysis, presentation & interpretation tools	0.9	1.7	2.6	5.2	Intuitive user interfaces for targeted user groups (1.7); analytical pipelines (1.3); data visualization & interpretation (0.9); tools to integration information across species (0.9); co-analysis of geo-spatial & genotypic data (0.4)
		Assistance & support	2.6	3.1	3.5	9.2	Data-management support and best practices (12); hosting of data (2.2); data-analysis know-how & manpower (1.7)
Community & networking	19.2	Synergies from working together	3.9	0.4	6.1	10.4	Cross-crop collective thinking and sharing of experiences & lessons learned, including with other sectors (4.4); new connections & collaborations with experts and similarly minded colleagues (2.6); landscape of projects & expertise in key areas such as information management (2.2); importance of minor crops and opportunity to benefit from bigger ones (0.9); cross-species analyses (0.4); identification of most pressing common needs, particularly in the area of information-management (0.4); participatory collaboration framework (0.4)

[1] Percentages were calculated using the total number of references to individual topics/ideas (229) as a denominator

Category	Fitting topics mentioned (% hits)[1]	Subcategories	How can DivSeek help you? (% hits)[1]	How can you help DivSeek? (% hits)[1]	Three top priorities for DivSeek? (% hits)[1]	Sum of three questions (% hits)[1]	Topics mentioned (in descending order of % hits)[1]
		Opportunities from joint action or advocacy	12	2.6	0.9	15.5	Increased visibility of own projects, expertise, data or tools (3.1); links to other stakeholders, organizations, initiatives or consortia of relevance, including on an international scale (breeders, informatics consortia, large IT companies, crop-specific consortia) (3.1); DivSeek "endorsement" for new projects (0.9); joint awareness raising and advocacy (0.9)
Research strategies	15.7	Accession-sampling strategies	0	1.3	1.7	3.0	How to sample individual accessions for genotyping/sequencing (single seeds, SSD, pools, etc.) and how to link back to original accessions (2.2); large-scale phenotyping strategies (0.4); cross-institutional characterization & evaluation of accessions (0.4)
		Genotyping & sequencing methods	0.9	0.9	1.7	3.5	Standardization of genotyping/sequencing methods (1.3); Genotyping/sequencing methods for small crops (0.9); Data quality of genotyping/sequencing methods and potentially confounding effects of epigenetics (0.9); Use of multiple reference genome sequences (0.4)
		Phenotyping methods	0	2.2	1.3	3.5	Standards, methods & experimental design (1.3); HTP field-based methods & environmental monitoring (1.3); HTP controlled-environment methods & platforms (0.9)
		Intended impacts	1.7	0.9	3.1	5.7	Mobilization of crop diversity for breeding (1.7); define and deliver a few pilot examples, including for a small crop (1.7); operational efficiency of genebanks (0.9); reduce duplication across genebanks (0.9); more targeted access to germplasm (0.4)
PGR-related rights	5.2	Data & germplasm-related rights & sharing principles	1.7	0	3.5	5.2	Data-sharing rules (2.2); germplasm access/sharing (non-Annex-1 crops, exchange with national genebanks, transfer of new materials to international genebanks) (1.7); align IP principles with other initiatives (0.9); effect of international treaties on germplasm access/exchange) (0.4)
Funding for projects	3.5	Access to funding	1.7	0.4	1.3	3.4	Greater leverage through joint fundraising efforts (1.7); DivSeek to provide funding for crop-specific projects (0.9); DivSeek adds value to institutional investments (0.4); in-kind contribution to DivSeek efforts (0.4)
Crop data sets	3.5	Data sets from crop-specific projects	0	3.1	0.4	3.5	Data sets from crop-specific projects (3.1); digitalization of valuable paper records (0.4)
Capacity-building	2.6	Tools and know-how in new technologies	0.9	0	1.7	2.6	Training in (tools for) data analysis (1.3); targeted at developing countries (0.9); training in new genotyping/phenotyping technologies (0.4)