

Bretting, Peter

From: Emily Marden [REDACTED]
Sent: Wednesday, December 09, 2015 8:23 AM
To: Bretting, Peter
Subject: Fwd: Notes from Dec 8
Attachments: Divseek Dec 8 Meeting Notes pp (1).docx

Hi Peter

Thanks for your note. I completely agree with your positions as did the others. I am forwarding the notes that Peter P and I took from the first part of the meeting. The governance discussion went well. More on that in notes to come.

Best

Emily

----- Forwarded message -----

From: Emily Marden [REDACTED]
Date: Wednesday, December 9, 2015
Subject: Notes from Dec 8
To: Susan McCouch <srm4@cornell.edu>, Susan McCouch [REDACTED]
Cc: Peter Phillips <peter.phillips@usask.ca>

Hi Susan,

Please see attached notes from yesterday morning's meeting. Peter and I have amalgamated our notes into this document.

Best,

Emily

Updates since Last Meeting

GPC:

Landscape projects identified. Of relevance to DS.

- Many projects have some DS elements but not exclusively DS; determination of what is DS component will evolve; SC asked for insights
- Projects being mapped geographically; these are self-identified and identified by Ruth, but there are many more. Could populate more.
- Can add questions and sorting mechanisms that could be used to identify and characterize projects.
- Other relational mapping possible (Phillips will investigate).
- Lacking developing world projects. Are they missed? Do they exist? Susan noted projects in India, for example, but may not want to be included on the list. Can we acknowledge them without causing difficulties? Do we create dialog with China or India in terms of internal projects being shared.

TREATY:

GB recognized DS. Have program of work for GLIS, as documented.

Bit of hard talking, but outcome extremely positive and enabling.

ASU study discussed and preliminary finding presented. Mostly unanalyzed case study descriptions at this point. SC will need a draft report with summary conclusions before it can consider how to use. After further discussions (later in the meeting), it was agreed by SC that while the report is needed, it would be premature to plan to use it as a discussion item at the January workshop. The SC awaits the draft final report, as planned, at the end of January.

GCDT

Have recognized concerns about what DS is going to do. Trust concerned about where heading. Currently DS is being pulled into policy domain, particularly at the Treaty's request. The Trust instead hopes DS can become a science platform, at the intersection of a triangle of genebanks, breeding programs, and the genomics/big data community to make accessions more useful. The Trust's core interest is in the genebank space. The Trust remains committed DS and is happy both to host and continue Peter's time for now.

CGIAR Consortium

Genetic resources are at the core of new CG portfolio for 2017. The question is how does DS connect into new portfolio? Portfolio will be focused on food systems and sustainability. Two platforms are proposed – genebanks and genetic gain. Need for connectivity on platforms is important. This could be an important role for DS.

Securing funding for genebanks going forward in current climate is important. All must do more with less. US\$90 million secured for genebanks; this is less than optimal, requiring clear performance indicators and management. Genetic gain platform will use high throughput genomics and related genome knowledge to increase impact in farmers' fields,

DivSeek could be a potential umbrella organization. DS must balance the perspective of supply driven with demand driven needs. We know that sequencing and resequencing is happening but not clear how that will shape food systems.

Within the CG Consortium, major changes are underway.

Treaty

Indicated clearly the Treaty does not want to assume lead/executive role. Rather prefers to contribute to enabling policy and governance, complemented with training and capacity development. Also exploring sequencing services by connecting partners--especially in countries where genomics sequencing is not available.

Happy to have lead/executive role at Trust or outsourced with third party (as with GLIS). This would require consultation at JFU and transition planning. The Trust and Treaty have not discussed how to take the Director model forward. Whatever chosen, with consultation, we can move to the new model.

Further Discussion:

- GPC agreed DS needs single person as point. Only way to make progress. Would still be involved. We would not be host but happy to continue doing certain activities (e.g. landscaping).
- Trust agreed an executive and operational team is a good way to work. This could advance drafting business plan to seek funding. From a trust perspective a key concern is identifying and assembling the skills needed at genebanks – there is a need for managers/capacity building and marrying fields of germplasm and genomics. This might involve stitching genebanks together in various arrangements.
- Consortium noted the need for transparency. As a final observation, the Consortium noted that big data is not getting sufficient importance—DS could help address that.

The SC went into Camera at this point of the agenda.

SC discussion about the Goal of DivSeek:

Susan: Are we an organization that brings people together or does DS actually want to accomplish something with the data (integration or dissemination). Could be both? Initially, DS was coordination. Since then, many people think it can be and should be more.

Sara: Asked how do we measure impact and demonstrate what we are doing. In this one year, what can we show? How have we and are we are going to be adding value. Measuring coordination is hard.

Dave Marshall: Suggested DS should undertake work with exemplar projects to illustrate options and strategies. Exemplars can cover the range of diversity of crops, resources and technical challenges. Yet another rice project is not sufficient. Apart from exemplar projects, DS could also be

- a technical advice/best practice broker.
- provide information hosting. Many crops struggle with this. Much of the hosting is organization based rather than crop based; this leads to a multiplicity of platforms, such as for wheat. Finding a framework to pool info is challenging; a major need
- offer advice on governance and legal framework
- to some extent it could be a funding broker.

All: While there might be an apparent tension with Treaty, in reality the Treaty focuses on the conceptual issue whereas DC focuses on the genetic issue. The Treaty is building links to other systems.

Ruaraidh: GLIS offers an good information platform, but more as an index than an information system. A key aspect of any information system is the unique identifiers. The PUIDs (DOIs) are linked to specific packets of seeds originating in the genebanks, which then links to all passport data. While used extensively for outgoing materials, the PUIDs are expensive to assign (est. \$1000/accession) and are not used as widely as desirable (e.g. not by those working with other materials and often not for derived varieties.

While some convergence is happening, in many species there are multiple annotation systems. It is healthy in a community to have multiple sequences – it is not terrible; it is the progression at the moment.

What is in GLIS? Anything to do with plant genomic resources for food and agriculture. Farm trials. Key elements are: when send material with SMTA, must make non confidential data available, and this will be included. Second, on the side of the recipient, they have an obligation to send back the results of their non confidential studies. Key is to provide the mechanism that allows them to report.

Susan: The GLIS concept is consistent with DivSeek. There is NO issue with this concept.

Andreas: WRT the mission of DivSeek and the relationship with the Treaty, he noticed resolution 3/2015. He strongly recommends the SC NOT accept the invitation. His view is DS does not have jurisdiction to opine. We should be enabling synergies as in the first bullet.

Question raised: How does the private sector fit into all of this; sorting this out and engaging would really distinguish DS from the Treaty.

We are not defined by the Treaty but are harmonious with it. We need to be inclusive of ALL relevant communities and not exclude anyone, whether private sector or non-members of the Treaty. Three groups currently have gaps: ventures that are funded; the non treaty members; private sector. Working with them would make DS very different.

DivSeek should work on structuring info that relates to genetic entities, rather than GLIS, which is pointers to information. One possibility is DS creates repository where people can put data. But, the challenge of developing long term infrastructure is that it would then need long-term institutional support. Reality is that it is easy to get money to set things up but to keep them going over the long term, hard to do.

Could DivSeek just brand/quality assure the product/system, with participants largely going their own ways but staying consistent in their coding and disclosure.

Elizabeth: Could DivSeek be an information platform to bring people together to find solutions. Looking at Ruth's list, putting all projects together, do they need a place to put data or do they need to find other data or ways to use data?. So perhaps for all those projects we can find a way for them all to collaborate by providing technical solutions and best practices.

Solving problems can be an intractable role to take. But looking at the landscape list, we can see many in the same field (e.g. maize), likely many also trying to solve similar problems. Multiple project may have a common challenge; bringing them together could help solve the isolation now plaguing projects. IN this context, could DivSeek be the matchmaker; either we take the initiative, looking down the list and getting groups together; or, someone looking help (e.g. bioinformatics expert) could trigger a matching process. One useful goal would be to encourage collaboration; people often have money but no practical expertise. Mission: "Bring people together"?

PUID is initially just for use with Treaty material but eventually it could (should?) be used for everything. * This could be part of the workplan. One goal could be to assign every program a PUID. Challenges/options include:

- costs to get PUID. We have to have a way to support an expense or to lower the cost of getting a PUID

- Encouraging granting agencies and journals to require PUIDs as norm for publishing/granting.
- Do we need a database to track everything once it is out?
- Could get involved in finding and giving out number.

Susan: Information is not currently hosted and combined in a meaningful way. Not integrated. Could use GOBI to integrate. But there are huge challenges for data integration. Right now data is diversifying faster than it is coming together. Right now projects just as often regenerate data rather than access and share existing data. Even more important for phenotyping, as regenerating phenotype results is usually not possible. But putting phenotypic data into a repository is more complex; people want both digested/summarized results and the statistics indicating probabilities.

Data handling data storage.

Andreas: How can we valorize the data? We could start to interact with people in terms of use cases. Have workshop to think about how to use all the data, i.e. the exemplary projects.

- How can we help genebank collections: management tools to become more efficient? Making collection more efficient by removing duplicates. Addressing collection management issues.
- Pre breeding; how do you move stuff from collections to purified lines and into use in breeding systems? Many mobilization issues. This tailors into breeding.

Rauri: Suggests efficiency from removing duplications not likely all that significant; fewer duplicates than thought; and cost of removing may be greater than maintenance costs for many species

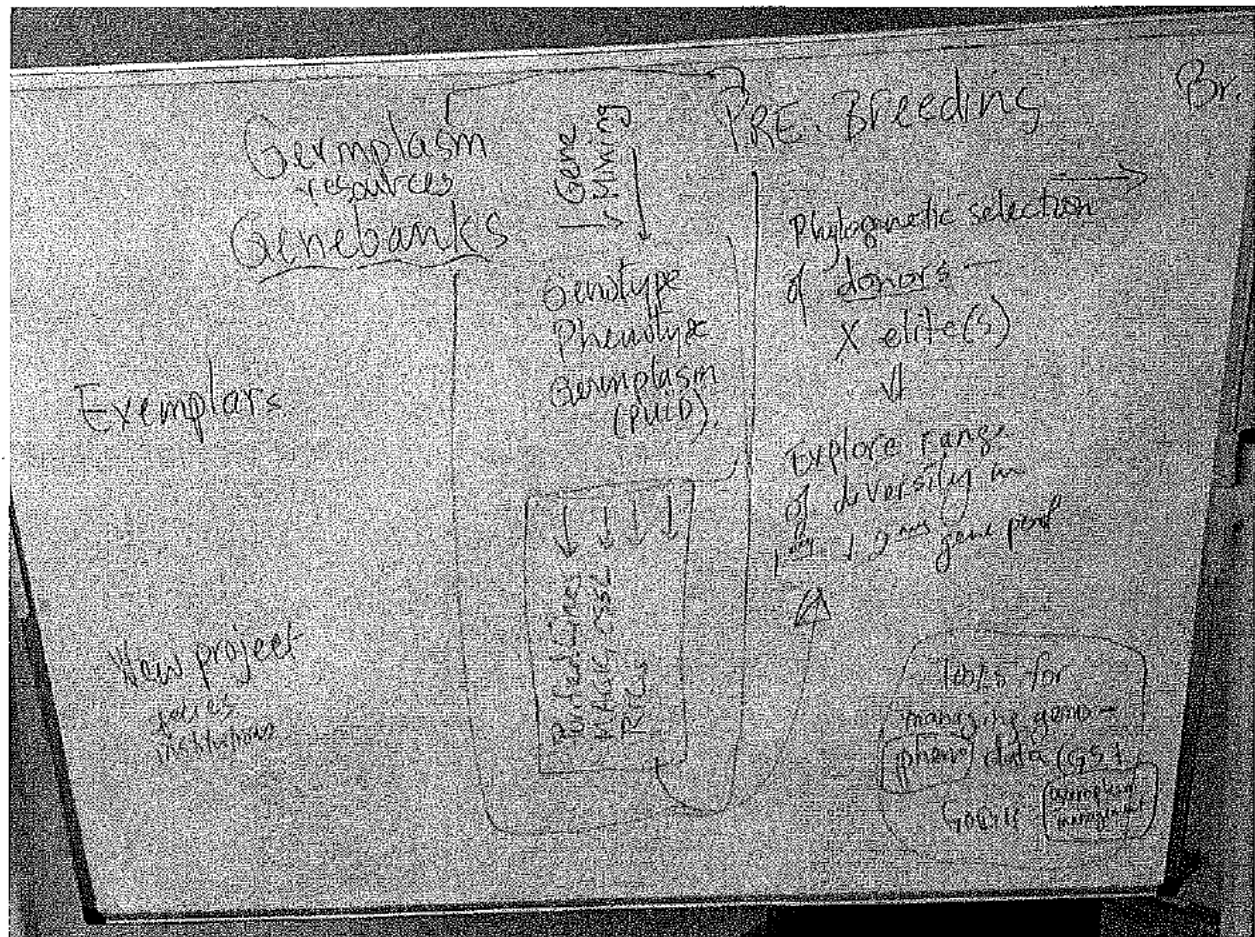
From this discussion the idea of a matrix, moving things from pre breeding to elite was discussed and mapped. Then an info management matrix cuts across the prebreeding space.

Right now a huge need to encourage pre breeding domain vis a vis genebanks. Getting things out of genebanks through prebreedings.

Big Idea: Do flagship project building on GOBI project.

Peter P: What would Canada want: explore wheat, pulse, other Canadian crops. Barley oats etc. Canada is doing a CFREF project linking genotype, phenotype and imaging through bioinformatics.

Crop wild relatives discussed. Could be fundable. Not much done by anybody. Breeders interested. Could relate well to sustainability/ climate change as well as food security platforms and priorities.



Notes/options related to Susan's visual

1. germplasm characterization and management: Can affect genebank management [germplasm management, after further discussion, decision that this is not probably a primary goal at this point. It will be a byproduct of other things, not the primary objective]
2. allele mining across a range of crop wild relatives, purified lines, magic CSSLs and RILs.
3. prebreeding design: Phylogenetic selection of donors x elites to explore range of diversity in primary and secondary gene pools

Potential summary of where we are:

There are array of options for a work plan:

Goals: support the development of biodiversity informatics to help genebank management, informed access, valorization

Activities: identify relevant players [we have project inventory and landscape] networking, interaction with funders and decision makers

1. Pulling together people to create and advance norms of good germplasm management to advance breeding

- ** integration of PUIDs is a core mission. Communicating with authors, publishers, also the GPC. [norms building]
- Promoting norms through workshops etc

	Norm	Limits	Action
Genetic materials in genebanks	PUIDs linked to seed packets and passport info	Only outgoing Treaty material now	Work to lower cost and get embedded in standards for journals and grants
Phenomics	Ontologies; some disclosure but limited repositories	Various ontologies that don't converge	Work to encourage more concordances among ontologies and repositories
Genomics	Multiple sequences, not all on common methods; mostly public repositories	Religious wars about which sequence system to use	Promote more common sequencing and disclosure of sequence model
Imaging	DOIs emerging as code;	few images in repositories; few digitized for further access	Set norms and promote
Publications	DOIs for journals but not for all pubs	Not universally used in follow-on publication	Promote as norm/standard for publication in this area
Algorithms	No standards for disclosure or sharing	No practice	Create pooling among public sector teams
Best practices	No standards for documenting or exchanging	No practice	Create model for codifying and sharing among public sector teams.

2. Capacity building: workshops on common issues; bring together projects by species or trait or technology to identify gaps and new options; knowledge transfer

- Deliverable: white papers recommendation, workshops,

3. Projects – anything from building, funding, managing, doing. We could initiate OR other could initiate and we could assist. There is a large menu. .

- **Needs to go with a glamorous project. Say \$10 million for GOBII for genebanks. Good project, but would take a significant funder. Does the gene mining and helps you manage genotype, phenotype and germplasm. Would allow you to take allele mining and trace through a breeding pedigree and see where it had ever been deployed. Allow you to find markers that are specific. Could make people choose partner to train teams.

Impacts on Governance

Agreed JFU needs an executive lead

Whether DS remains with Trust or moves on to partner with another entity will be determined by the workplan. If a move is appropriate in response to the workplan, LOIs should be solicited for 3-5 leadership.

DS JFU needs to find a way to move from the consensus model to allow partners to take lead on areas of greatest interest/capacity and to not be beholden for activities not within their mandate but appropriate for DS.