3-2018

Plant genetic resources for food and agriculture: opportunities and challenges emerging from the science and information technology revolution

Michael Halewood
Bioversity International, Rome, Italy

Tinashe Chiurugwi
NIAB

Ruaraidh Sackville Hamilton
International Rice Research Institute, Manila, Philippines

Brad Kurtz

Emily Marden
University of British Columbia

Follow this and additional works at: https://lib.dr.iastate.edu/soc_las_pubs

See next page for additional authors

Part of the Agricultural Science Commons, Biotechnology Commons, Plant Breeding and Genetics Commons, and the Science and Technology Policy Commons

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/soc_las_pubs/28. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.

Plant genetic resources for food and agriculture: opportunities and challenges emerging from the science and information technology revolution

Abstract

Over the last decade, there has been an ongoing revolution in the exploration, manipulation, and synthesis of biological systems, through the development of new technologies that generate, analyse and exploit big data. Users of Plant Genetic Resources (PGR) can potentially leverage these capacities to significantly increase the efficiency and effectiveness of their efforts to conserve, discover and utilise novel qualities in PGR, and help achieve the Sustainable Development Goals (SDGs). This review advances the discussion on these emerging opportunities and discusses how taking advantage of them will require data integration and synthesis across disciplinary, organisational and international boundaries, and the formation of multi-disciplinary, international partnerships. We explore some of the institutional and policy challenges that these efforts will face, particularly how these new technologies may influence the structure and role of research for sustainable development, ownership of resources, and access and benefit sharing. We discuss potential responses to political and institutional challenges, ranging from options for enhanced structure and governance of research discovery platforms to internationally brokered benefit-sharing agreements, and identify a set of broad principles that could guide the global community as it seeks or considers solutions.

Disciplines

Agricultural Science | Biotechnology | Plant Breeding and Genetics | Science and Technology Policy

Authors

Michael Halewood, Tinashe Chiurugwi, Ruaraidh Sackville Hamilton, Brad Kurtz, Emily Marden, Eric Welch, Frank Michiels, Javad Mozafari, Muhamad Sabran, Nicola Patron, Paul Kersey, Ruth Bastow, Shawn Dorius, Sonia Dias, Susan McCouch, and Wayne Powell

1   **Plant genetic resources for food and agriculture: opportunities and challenges emerging from the**

2   **science and information technology revolution**

3   Michael Halewood[1], Tinashe Chiurugwi[2], Ruaraidh Sackville Hamilton[3], Brad Kurtz[4], Emily Marden[5], Eric

4   Welch[6], Frank Michiels[7], Javad Mozafari[8], Muhamad Sabran[9], Nicola Patron[10], Paul Kersey[11], Ruth

5   Bastow[12], Shawn Dorius[13], Sonia Dias[14], Susan McCouch[15], and Wayne Powell[16]

6   **Author Affiliations:**

7   1. Bioversity International, Via dei Tre Denari, 472/a, 00054 Maccarese, Rome, Italy

8   2. NIAB, Huntingdon Road, Cambridge, CB3 0LE UK

9   3. T. T. Chang Genetic Resources Center, International Rice Research Institute, DAPO Box 7777, Metro

10   Manila, Philippines

11   4. Independent Crop Biodiversity and Intellectual Property Expert, 25057 River Ridge Road, Adel, Iowa

12   50003, USA

13   5. University of British Columbia, Peter A. Allard School of Law, 1822 East Mall, Vancouver, BC Canada

14   V6T 1Z1

15   6. School of Public Affairs, College of Public Programs, Arizona State University, 411 North Central

16   Avenue, Suite 463, Phoenix, Arizona 85004-0687, USA

17   7. Independent Crop Biodiversity and Intellectual Property Expert, Technologiepark 38, 9052 Gent,

18   Belgium

19   8. Agricultural Research, Education and Extension Organization, Yemen St, Chamran Freeway, Tehran,

20   Iran

21   9. Indonesian Centre for Biotechnology and Genetic Resources, jl tentara pelajar No 3A, kampus

22   penelitian pertanian Cimanggu, Bogor 16111, Indonesia

23   10. Earlham Institute, Norwich Research Park, Norwich, Norfolk, NR4 7UH, UK

24   11. EMBL-The European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire,

25   CB10 1SD, UK

26   12. Global Plant Council, Bow House, 1a Bow Lane, London, EC4M 9EE, UK

27   13. Department of Sociology, Iowa State University. 308 East Hall, Ames, IA 50010, USA

28   14. Secretariat of International Treaty on Plant Genetic Resources for Food and Agriculture, Viale delle

29   Terme di Caracalla, 00153 Rome, Italy

30   15. Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Cornell

31   University, 240 Emerson Hall, Ithaca, NY 14853, USA

32   16.   SRUC (Scotland's Rural College), Peter Wilson Building, West Mains Road, Edinburgh, EH9 3JG, UK

33   **Corresponding Author:** Wayne Powell, Email: <u>wayne.powell@sruc.ac.uk</u>, Phone: +44 1315 354001

34   **Brief heading:** Emerging opportunities and challenges for plant genetic resources for food and

35   agriculture

36   **Lab-associated Twitter handle:** @SRUC

37   **Total word count** for the main body of the text: 8300

38   **Number of figures**: 2 (both to be published in colour)

Summary

40   Over the last decade, there has been an ongoing revolution in the exploration, manipulation, and

41   synthesis of biological systems, through the development of new technologies that generate, analyse

42   and exploit big data. Users of Plant Genetic Resources (PGR) can potentially leverage these capacities to

43   significantly increase the efficiency and effectiveness of their efforts to conserve, discover and utilise

44   novel qualities in PGR, and help achieve the Sustainable Development Goals (SDGs). This review

45   advances the discussion on these emerging opportunities and discusses how taking advantage of them

46   will require data integration and synthesis across disciplinary, organisational and international

47   boundaries, and the formation of multi-disciplinary, international partnerships. We explore some of the

48   institutional and policy challenges that these efforts will face, particularly how these new technologies

49   may influence the structure and role of research for sustainable development, ownership of resources,

50   and access and benefit sharing. We discuss potential responses to political and institutional challenges,

51   ranging from options for enhanced structure and governance of research discovery platforms to

52   internationally brokered benefit-sharing agreements, and identify a set of broad principles that could

53   guide the global community as it seeks or considers solutions.

54   **Keywords:** Access and benefit sharing, Big data, Data integration, Farmer's rights, Global governance,

55   Plant genetic resources for food and agriculture (PGRFA), Sustainable development goals (SDGs),

56   Synthetic biology,

1. Introduction

58   Technologies for generating and analysing large quantities of genotypic, phenotypic and environmental

59   data are evolving at accelerating rates; so too are technologies and methods for synthesising genetic

60   materials (Policy Science for Environment, 2016). New technologies for high-throughput assays and

61   synthesis of genetic materials are revolutionising biology. The development of techniques for highly

62   parallel, genomic sequencing has been followed by other methods for measuring the current molecular

63   state of cells and organisms, for predicting classical phenotypes in an automated manner (Furbank &

64   Tester, 2011), and even for re-engineering the content and function of living systems (Noman *et al.*,

65   2016). These technologies have led to the rapid generation of large amounts of data describing

66   biological systems, and the analysis and interpretation of these data using statistical and computational

67   expertise. These changes have transformed biology into an information-rich science, where the

68   integration and interpretation of large quantities of data informs both the design and nature of new

69   hypotheses and the application of existing results. At the same time, they are raising questions about

70   the applicability of governance regimes for plant genetic resources for food and agriculture (PGRFA) that

71   were established primarily to deal with the exchange of genetic resources.

72   This review provides an overview of how plant gene banks, plant breeders, national programmes and

73   farmers can potentially take advantage of these capacities to significantly increase the efficiency and

74   effectiveness of their efforts to discover, conserve, and use new qualities in plant genetic resources.

75   As part of this enterprise, a range of actors will need to generate, access, integrate and synthesise data

76   that is widely dispersed across organisational and international boundaries, and work through

77   international partnerships that bring together complex portfolios of skills, sources of information, and

78   perspectives. To be successful, these partnerships will need to overcome a number of institutional and

79   policy challenges that might otherwise affect the willingness of partners to cooperate in research and

80   development activities, including generating and sharing digital genetic and traitdata describing plant

81   genetic resources (PGR). One such challenge is revising structures that currently incentivise people to

82   treat data and information confidentially (e.g., due to competition for monetary rewards, intellectual

83   property rights, grants, and publications), sharing it only among trusted friends or colleagues, or under

84   restrictive licences. Another challenge concerns contested claims about how benefits derived from the

85   use of data should be shared among those that have contributed to the evolution and conservation of

86   those resources. A range of collaborations and open-access data repositories have recently been

87   criticised for facilitating uses of genomic data in ways that allow commercial users to take advantage of

88   genetic resources without having to share benefits as prescribed by international access and benefit-

89   sharing laws (Hammond, 2016; The International Civil Society Working Group on Synthetic Biology,

90   2016). The laws in question require users to share monetary benefits in exchange for access to, and use

91   of, material genetic resources in the development of new commercial products; they do not extend in

92   scope to genetic sequence information. A growing number of developing countries and civil society

93   organisations are calling on the United Nations to address this issue by developing new international

94   obligations to share benefits derived from the use of genetic sequence data.

95   This review examines options for addressing these challenges through enhanced governance

96   arrangements. This is timely in part because the 'omics' revolution in the biological sciences has

97   considerable disruptive potential for changing the flows of information, the nature of partnerships, and

98   the range of products and benefits that can be generated through PGR conservation and applied plant

99   breeding. Meanwhile, the policy environment has lagged behind, with a concomitant shortage of shared

100   norms to guide the resolution of contested claims related to how omics-driven research and

101   development is conducted in relation to PGR. The kinds of governance arrangements eventually put in

102   place, and the manner in which contested benefit-sharing claims are resolved, could have considerable

103   impact on the way in which research and development in the agricultural sector is perceived and

104   conducted.

105   **2. Technological advances and their utility for gene banks and breeding, and**

106   **longer-term contributions to SDGs**

107   *The impact of genetic technologies on germplasm repositories and seed banks*

108   Historically, genetic variation was recognised and tracked based on visual assessment of phenotypic

109   variation, but since the development of molecular marker technology in the 1980s, it can be assayed

110   directly at the DNA level. There are many genotyping platforms (Goodwin *et al.*, 2016), but all are

111   designed to do essentially the same thing: identify differences in the genetic sequences of individuals,

112   and record the differences (polymorphisms) and monitor their presence or absence in specific

113   individuals in a systematic way, often using a reference genome coordinate system. In the context of

114   germplasm repositories (here we refer to all collections held by private individuals, companies, national

115   or international bodies), a.k.a. "gene banks", this has a number of obvious applications. Firstly, it makes

116   possible the extensive characterisation and traceability of the stocks that are currently held by gene

117   banks. Genomic analysis allows the level of variation among individuals in a single seed pack, gene bank

118   accession, or a number of accessions to be ascertained and accurately quantified. This can better inform

119   the quality control, maintenance, distribution, and use of gene bank stocks. Moreover, extensive

120   genotyping, linked to measured traits, allows repositories to be searched for strains containing desired

121   genetic elements and/or trait characteristics, and the production of new products that improve the

122   sustainability, diversity, and resilience of crop plants, which is particularly important given the urgent

123   need to feed and nourish a growing global population in the face of substantial change to the natural

124   environment.

125   ***Integrating big data into breeding programs***

126   Generating sequence information *per se* is no longer a bottleneck for crop improvement. Phenotypic

127   characterisation has historically been more problematic, but increasingly, molecular phenotypes (e.g.

128   gene expression and ion content) can be used as indicators of physiological or performance phenotypes,

129   while quantitative imaging techniques using remote sensing can directly measure plant architectural and

130   stress response characteristics in a variety of experimental set-ups (e.g. automated greenhouse and field

131   settings under drone surveillance). Statistical models can be constructed to predict the breeding value of

132   an individual, given its genomic composition (genomic selection); and an optimal breeding scheme can

133   be designed in the light of such predictions. Genomic selection has been associated with major

134   performance gains in livestock species, and it has similar potential in plant species. Accurate prediction

135   is enhanced not only by access to *more* data, but to more variation in the data – which means sampling

136   the gene pool as extensively as possible, and assaying under the widest range of environmental

137   conditions. If information can be collected, it is possible that the method will develop sufficient power to

138   accurately predict G x E (Genotype x Environment) interactions, allowing for the development of specific

139   crops tailored to particular environmental conditions. Dedeurwaerdere (2013, p. 369) notes that, the

140   "information technology revolution has dramatically expanded the possibilities of distributed

141   coordination …" in the use of genetic resources. Indeed, the increasing generation and use of big data by

142   farmers themselves (both as inputs into and outputs generated by agronomic decisions) could

143   potentially create a huge reservoir of knowledge about plant performance (including stress tolerance,

144   nutritional quality and overall yield) in a far wider range of climates, soils, and management regimes

145   than could be tested by a single breeder, research team, or organisation (Satizábal *et al.*, 2012; van

146   Etten *et al.*, 2016, 2017). This information, if made available to breeders and biological engineers, has

147   great potential to feedback into further improvement programmes. A more formal and extensive

148 partnership between farmers, researchers and other actors to facilitate the flow of information stands

149 to substantially enhance benefits to the variety of plant genetic resources stakeholders.

150 **Technologies for identifying and creating genetic variation**

151 *Molecular marker (genomics)-assisted germplasm curation, research and breeding*

152 All crop improvement practices aim to capture (within elite lines) genetic variants that confer desirable

153 traits. The ability to accurately identify and track genome-wide genetic variation or individual molecular

154 variants across generations of individuals offers a powerful tool for germplasm managers, basic

155 researchers, and plant breeders (Collard & Mackill, 2008; McCouch *et al.*, 2012). For example, gene bank

156 managers utilise molecular markers to establish and validate the identity of accessions in their

157 collections, to determine genetic relationships among individuals, to perform gap analysis to guide

158 collecting efforts, and for allele mining to identify accessions that carry particular alleles (traits) of

159 interest. Basic biological researchers use genomic and other "omics" analyses to characterise the

160 structure, function, and evolutionary significance of genes and alleles, to study plant development and

161 response to environment, and to understand speciation and the implications of diversity at the

162 individual, population and ecosystem levels. Applied breeding programs use molecular marker data to

163 identify parents for crossing, to select offspring carrying favourable or deleterious alleles in segregating

164 populations, and to perform genomic prediction.

165 *Genetic and genome engineering*

166 Since the 1980s it has been possible to randomly insert new genetic material into the genomes of plants.

167 The first genetically modified crops to be commercialised were tomatoes with extended shelf life (1994),

168 insect resistant potatoes (1995), herbicide (glyphosate) resistant soy (1996) and virus resistant papayas

169 (1998). Genetically modified crops are now grown on 181.5 million hectares of land, by 18 million

170 farmers (Stevenson *et al.*, 2013; James, 2014). In the past decade or so, emerging technologies (such as

171 programmable nucleases, e.g. zinc finger nucleases and RNA-guided Cas9 [i.e. CRISPR-associated protein

172 9] from bacterial CRISPR systems) have enabled so-called precision genome engineering (or genome

173 editing): the induction of targeted modifications to the genome, its contexts (e.g. epigenetic marks) or

174 its outputs (e.g. transcripts) (Schiml & Puchta, 2016; Petolino *et al.*, 2016). Targeted genome

175 modifications include the induction of mutations at pre-selected loci to disrupt the function of one or

176 more specific genes; the editing of existing sequences to reproduce ancient alleles or to introduce novel

177 alleles; or the introduction of new genetic material into specific loci or regions of the genome. It is also

178   possible to change DNA modifications, such as methylation, in order to modulate gene expression.

179   When coupled with the ability to chemically synthesise DNA molecules at ever diminishing costs,

180   genome engineering may enable multiple novel variations to be designed and tested at any desired

181   genetic locus, including in multifactorial combinations (Puchta, 2017).

182   Although some of these technologies are still inefficient and difficult to execute, they are being

183   developed for numerous food and non-food crops, and progress continues apace. Genome engineering

184   and synthetic biology technologies have the potential to vastly reduce the time taken for knowledge

185   generated in the laboratory to transition into marketable products by allowing the direct introduction of

186   favourable alleles into agronomically valuable germplasm, thus reducing the number of breeding cycles

187   required. One day, this may mean that it is possible for a farmer to request that a targeted set of

188   changes be made to a highly valued cultivar as part of the breeding process. In response, a new genetic

189   trait or combination of genetic characteristics could be rapidly designed and introduced into a cultivar to

190   improve its resilience to stress, nutritional quality or architectural characteristics, making it a better fit

191   for either the traditional cropping system or the modern agricultural landscape. These applications have

192   the potential to bypass the direct use of specific *physical* host plants that may have played an important

193   role in identifying the novel traits, with wide ranging implications for owners, managers and users of

194   PGR.

195   ***Information technologies underpinning plant genetic resources***

196   Information technology has played as important a role as genomic technology in the evolution of crop

197   improvement strategies. The ongoing improvement in the performance of computers, driven by the

198   ever-increasing miniaturisation of transistors, was noted by Moore as long ago as 1965 (Moore, 1965).

199   Subsequently, the development of many other aspects of computer hardware (parallelisation, storage,

200   networking, etc.) has enabled, and also been driven by, the data revolution in almost every field of

201   study. In particular, the development of the Internet, the World Wide Web, and the explosive growth in

202   mobile communications networks, have put much of this computing power and the data that lies behind

203   it in the hands of citizens worldwide, including those interested in PGR (scientists, breeders, farmers,

204   and consumers). However, much of the relevant data is highly dispersed, has limited compatibility, and

205   is in practice hard to interpret except by specialists.

206   Apps are computer programmes that have been optimised for a particular purpose, and commonly used

207   on mobile computing devices. Compared with traditional tools, apps are often simple to use, and each

208   one developed to address a specific, limited, well-defined use case, frequently by opportunistic

209   entrepreneurs and social entrepreneurs (the barriers to entry in the app development market are

210   relatively low). Increasingly, such apps are in use by farmers, even in low and middle-income countries,

211   to source seeds and other agricultural inputs, improve agronomic and pest management practices, and

212   optimise market decisions. The provision of information about the genetics and performance of crop

213   varieties in particular environments through apps could assist farmers in the selection of varieties

214   appropriate to their conditions and cropping systems. Moreover, the collection, by farmers, of detailed

215   measurements of the actual environments in which crops are grown could enable the development of

216   more precise and sophisticated modelling of G x E interactions. The potential for such advances, while

217   holding tremendous promise, is still largely untapped.

218   **3. The challenges that must be overcome to realise emerging R&D**

219   **opportunities**

220   Access to large-scale sequence and phenotype information at unprecedented scales is providing new

221   opportunities to accelerate the application of basic research. This includes the ability to formulate

222   testable hypotheses about the genetic architecture of quantitative variation, the genes and biological

223   pathways involved, and the causal variants responsible for the inheritance of complex traits in diverse

224   species (Hamblin *et al.*, 2011; Lipka *et al.*, 2015; Sardos *et al.*, 2016; Yano *et al.*, 2016). It should be

225   noted that raw sequence information, if it is to be correctly interpreted and exploited, needs to be

226   integrated with an intimate knowledge of the biology of the species under consideration, the phenotype

227   or performance of the individuals or population that has been sequenced, and the agro-ecosystem in

228   which they have been grown, including the cultural context and farmers' management practices.

229   Further, when experiments are implemented appropriately, taking into account experimental and

230   mating design (Cavanagh *et al.*, 2008; Ersoz *et al.*, 2009), there are opportunities to unify discovery

231   biology with breeding. In other words, breeding programs can become 'test beds' for hypotheses about

232   G x E interaction as well as platforms for the development and deployment of new varieties (Poland,

233   2015). Realising this potential will require the ability to work at different scales, extending from

234   molecules to landscapes within a quantitative biology framework (Cooper *et al.*, 2014), and will require

235   greater collaboration between breeders, growers and the biological research community. A potential

236   model describing such a framework is shown in Figure 1.

237   **Figure 1.**

238   *Technical impediments to sharing, tracking, annotating, and linking data*

239   Some of the technical impediments that need to be overcome to facilitate data-integration and the

240   potential for data-sharing include: (1) data are fragmented and dispersed across organisations and

241   international borders and are not managed following the FAIR principles of Findability, Accessibility,

242   Interoperability, and Reusability (Wilkinson *et al.*, 2016); (2) inadequate systems exist for logging and

243   tracking PGR as well as metadata related to PGR, e.g. there is no universally agreed-upon system for

244   permanently and uniquely identifying PGR (e.g. publications do not provide traceable, permanent

245   unique identifiers for PGR); and (3) radically different approaches to data management and sharing

246   within and across public and private sectors due to fundamentally different objectives and low levels of

247   mutual trust.

248   These challenges are further complicated by the fact that a number of different kinds of data need to be

249   integrated, including genetic sequence, phenomic, environmental, and GIS (geographic information

250   system) data. Some have suggested that data on PGR in genebanks should be deposited in digital data

251   repositories (mirroring the physical repositories they describe), and made available, consistent with

252   national and international agreements, through standard application programming interfaces (API) to

253   scientists, breeders, farmers, and entrepreneurial tool developers. Several initiatives have been

254   established to promote this interoperability, including DivSeek (www.divseek.org), Global Open Data for

255   Agriculture and Nutrition (GODAN: www.godan.info), the Research Data Alliance (RDA: [www.rd-](http://www.rd-alliance.org)

256   [alliance.org](http://www.rd-alliance.org)), and the Breeding API (BrAPI: https://brapi.org). The Global Information System of the

257   International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) seeks to link existing

258   information systems on PGRFA. It has established a new mechanism to facilitate this by identifying

259   PGRFA using Digital Object Identifiers (DOIs: [www.fao.org/plant-treaty/areas-of-work/global-](http://www.fao.org/plant-treaty/areas-of-work/global-information-system/en)

260   [information-system/en](http://www.fao.org/plant-treaty/areas-of-work/global-information-system/en)). The goal is to promote the networking of high quality data repositories (and

261   data access channels) with efficiently designed data input systems. If successful, the deployment of DOIs

262   should increase the quantity and diversity of PGR-related data, reduce utilisation barriers to the plant

263   material conserved by germplasm repositories, and facilitate their use by growers, plant breeders, and

264   biological engineers. Adoption of standardised protocols would facilitate pooling of such data from

265   across organisations for comparative analyses and collaborative work, and could lower entry barriers

266   that currently limit farmer involvement in translational agricultural research and development.

267   Experience from other areas of biological research provides some clear models for how this need for

268   data integration can be addressed. The transformative effect of bioinformatics on many biological

269   questions owes much to open-access data. A key decision was made in the early 1980s to operate three

270   international repositories to hold nucleotide sequences, i.e. GenBank at the National Center for

271   Biotechnology Information (NCBI) in the USA, the European Molecular Biology Laboratory - EMBL Data

272   Library (today the European Nucleotide Archive), and the DNA Database of Japan (DDBJ) (Lawson &

273   Rourke, 2016). This unprecedented degree of collaboration among members of the international

274   research community paved the way for persuading the leading scientific journals to require data

275   publication as a precondition of publishing a traditional article. This new norm has resulted in large

276   quantities of nucleotide sequence data being made publicly available generally without any claims of

277   intellectual property passed on by the data providers or the database operators (see Notes S1). These

278   data now include the sequences of humans, other mammals, birds, fish, insects, microbes, and over 120,

279   000 flowering plant species. More recently, newer models for pre-publication of data and manuscripts

280   (deposition in advance of formal acceptance by a refereed journal) have been suggested, following the

281   norms of other communities (Authors: Toronto International Data Release Workshop, 2009). Further,

282   alliances of interested parties have formed to develop data models and appropriate structures for

283   interfacing between public and private data. One such example in the context of medical bioinformatics

284   is *The Global Alliance for Genomics and Health (http://genomicsandhealth.org)*.

285   ***Political and institutional impediments***

286   Bioinformatics capacities have not evolved in a political and institutional vacuum. They have gained

287   prominence during a period of considerable discord within the international community concerning the

288   proper balance of incentives for the development of advanced agricultural technologies (in the form of

289   intellectual property rights) on one hand, and for the conservation and sustainable use of genetic

290   resources (in the form of access and benefit sharing laws) on the other.

291   Until the late 1960s, plant genetic resources were generally treated as 'global public goods'. In the

292   decades that followed, technologically advanced countries pushed for international recognition of

293   intellectual property protection for living materials, through the UPOV (International **U**nion for the

294   **P**rotection **o**f New **V**arieties of Plants) Convention and the Uruguay Round of GATT (General Agreement

295   on Tariffs and Trade) negotiations. This led to disquiet, particularly among developing countries that

296   were the historic sources of much of the genetic diversity of the crops being commercialised and

297   protected. Developing countries pushed back through negotiations under UNEP leading to the

298   Convention on Biological Diversity 1993 (CBD: www.cbd.int). They asserted insisting on the recognition

299   of their sovereign rights to regulate access to genetic resources within their borders, with the

300   expectation of negotiating access and benefit sharing agreements with foreign access-seekers (or

301   'bioprospectors').

302   Discontent with the impact of the CBD on benefit sharing led to the negotiation of the Nagoya Protocol

303   on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their

304   Utilization (Nagoya Protocol), which came into force in 2014. Under the Nagoya Protocol, Contracting

305   Parties agree to put mechanisms in place for monitoring and enforcement of bilaterally negotiated

306   access and benefit sharing agreements. At the moment, it is too early to predict how the Nagoya

307   Protocol will impact stakeholders' willingness to share genetic resources for use in agricultural research.

308   Meanwhile, the 2004 International Treaty on Plant Genetic Resources for Food and Agriculture

309   (ITPGRFA) created a multilateral system of access and benefit-sharing (MLS) for Contracting Parties and

310   international organisations. The MLS provides facilitated access to the genetic diversity of 64 crops and

311   forages for the purposes of conservation and use for agricultural research, training and plant breeding.

312   Commercial users of material accessed under the multi-lateral system are obliged to make financial

313   payments to an international benefit-sharing fund under prescribed circumstances. To date, no

314   obligatory payments have been made under this system. In turn, some potential material providers are

315   not following through on commitments to make genetic resources available through the system. These

316   lapses led to the launch, in 2013, of a process for enhancing the multilateral system of access and

317   benefit sharing through renegotiation of its basic terms. If and when agreed, these new terms will be

318   reflected in a revised Standard Material Transfer Agreement (SMTA).

319   It is important to note that the CBD, its Nagoya Protocol, and the ITPGRFA link benefit-sharing

320   obligations to accessing and using physical *material* containing functional units of heredity (e.g. seeds

321   and cuttings). They do not specifically regulate access to digital data, e.g. genomic sequence or

322   phenotypic data (see Notes S2). All three agreements refer to digital research data as a potential benefit

323   to be shared in return for access to genetic resources. In fact, all three agreements were negotiated

324   without much discussion or debate about how ongoing technological breakthroughs (e.g. sequencing,

325   phenotyping, and bioinformatics) might eventually make it possible to take advantage of genetic

326   resources without the need to access the physical resources. Thus, even if the agreements were to

327   operate as intended, they would not directly address concerns that use of open-access sequence data

328   (and other related big data) will make it possible to profit from the use of genetic resources without

329   benefit-sharing obligations (see Notes S3).

330   In light of this recent history, it is perhaps not surprising that some country and regional representatives,

331   civil society and farmers' organisations have voiced concerns that technological breakthroughs in

332   genomic breeding, gene editing, and gene synthesis will widen the technology gap, and concomitant

333   economic disparities, between developed and developing countries. They worry that these

334   breakthroughs will exacerbate tensions associated with the unrealised expectations of monetary

335   benefits accruing from access and benefit sharing laws. These stakeholders note that at present the

336   requisite technological capacities principally reside in elite research institutions in the global North. They

337   are sceptical that these new capacities will be used to develop technologies targeted at resource-poor

338   farmers working in vulnerable agricultural systems. This has led to demands (by some civil society

339   organisations and developing countries) that research organisations stop providing unregulated open

340   access to genetic sequence data until benefit sharing issues can be addressed (Hammond, 2016; The

341   International Civil Society Working Group on Synthetic Biology, 2016). If efforts are not made to

342   enhance trust and inter-stakeholder cooperation, these controversies might ultimately undermine the

343   development of an open-science culture, slowing the rate of the scientific advance and crop

344   improvement. On the other hand, there are also reasons to view the development of legally binding

345   solutions at the level of the United Nations with caution. Reichman, Uhlir and Dedeurwaerdere (2016, p.

346   81) summarise the overriding institutional challenge:

347   "the resulting fears of … 'biopiracy' if left unchecked, [threaten] to destabilize the pre-existing

348   system of formal and informal exchange of both *ex situ* and *in situ* genetic resources on which …

349   agricultural research and applications have traditionally depended. By the same token,

350   overzealous regulatory measures to defend sovereign rights to these genetic resources could

351   perversely shut down that same system of exchanges, with potentially serious consequences for

352   global scientific research."

353

354   **4. Renewed governance structures for PGR (and related big data)**

355   *"In the future, if the multiple source [innovation] model is used, we will see a switch from the*

356   *transfer of institutional models and blueprints for research methods from centers to clients to an*

357   *approach where emphasis is on seeking out, understanding, and learning from innovators in*

358   *their local context. An implication for resource allocation in agricultural research will be a major*

359   *increase in the funds given to information exchange and networking activities whereby local*

360   *institutions are in direct contact with each other and "centres" are no longer seen as the hub. A*

361   *further implication of the model would be that increasing access to, and control of, these*

362   *networks would be in the hands of the poorer groups."*

363   Biggs (1990)

364   The forgoing analysis underscores the need for enhanced governance of the generation and use of

365   genetic sequence data and related information about PGR. This enhanced governance is necessary in

366   order to promote trust and transparency amongst different stakeholders, and encourage the

367   development and use of knowledge and technologies that ultimately advance the sustainable

368   development goals. The questions we examine in this section are: 'what forms of governance are

369   needed? What kinds of interventions would be necessary to enhance the overall operation of the

370   existing patchwork of organisations, institutions and practices in furtherance of the SDGs (see Figure 2)?'

371   On one hand the new bioinformatic capabilities risk exacerbating many of the unresolved issues related

372   to the governance of agricultural research generally and genetic resources more specifically.   On the

373   other hand, we are hopeful that, these new 'disruptive technologies' have the potential – if properly

374   governed – to transcend some of those long-standing tensions. This is partly because they make it

375   possible to engage a broader range of interested parties in the research and development process in

376   ways that are economically efficient, practical, and attuned to non-market considerations. For example,

377   they can radically lower the costs of local level needs assessments and facilitate crowd-sourced farmer

378   evaluation of materials across a broad range of agro-ecosystems.

379   Of course, these are not entirely new questions. Over the course of the last 30 years, there have been a

380   number of studies, from a range of theoretical perspectives, analysing the influence of organisational

381   structures on agricultural research and development in general, and more particularly on efforts to

382   conserve, add value to, share and exploit PGR. Many of those studies were conducted before the recent

383   increases in technological capacities outlined above; nonetheless, they can still provide useful insights.

384   ***Modular architecture for commons-based production***

385   Dedeurwaerdere (2013) considers a range of case studies of 'commons-based production' of public

386   goods in whole genome sequencing, wheat breeding, animal breeding programmes, and research on

387   root nodule bacteria for use in soy bean production. He concludes that many of the most successful

388   programs have been characterised by "modular architectures" which allow the pooling of input from

389   many individuals from diverse backgrounds, focus and geographical location. By extension, the

390   innovation systems through which these goods are produced necessarily require the inputs of many

391   actors, none of whom can act alone. Dedeurwaerdere notes a second common feature in successful

392   case studies: a critical density of the actors participating in the commons-based production efforts are

393   motivated by non-market incentives, for example, recognition by scientific peers, access to funding,

394   commitment to sustainable development, biodiversity conservation, or other broader social goals.

395   Furthermore, the goods produced through these modular architectures of commons-based production,

396   e.g. improved breeding lines and whole genome sequences, tend to be treated as pre-commercial. That

397   is not to say that information commons are exclusive of private interests or linkages to further

398   development of appropriable, commercial goods; indeed, such linkages frequently exist, but the goods

399   are not generated exclusively for this purpose.

400   Dedeurwaerdere's work fits within an emerging field of analysis of 'new commons' focusing on purely

401   culturally created goods (Madison *et al.*, 2010; Ostrom & Hess, 2010). It builds on Elinor Ostrom's

402   institutional analysis of factors contributing to the sustainable management of common-pool natural

403   resources (Walker *et al.*, 1990). Common-pool natural resources, e.g., forests and watersheds, are

404   defined as rivalrous (i.e. one person's use of a resource detracts from others' use of the same resource),

405   and non-excludable (i.e. it is difficult or impossible to prevent others from accessing the resource). The

406   natural resources and users studied by Ostrom were necessarily limited in number and geographic

407   space. On the other hand, the new cultural commons pertaining to PGR that are emerging as a result of

408   new information technologies and capacities may be distributed around the world, with potential to

409   dramatically scale-up the number of participants. In the case of digital information, one person's use

410   does not impinge on others' use of the same resource. Thus, information is non-rivalrous. It is also

411   virtually impossible to exclude others from accessing it. Perhaps the biggest difference between cultural

412   and natural resources commons is that the former have to be created through commons-based

413   production systems, while the latter already exist. As a result, the 'social dilemmas' to be addressed with

414   respect to both commons are very different. Madison *et al.* (2010) state that:

415   "…unlike resources in the natural world, resources of information and expression must be

416   created before they can be shared. Because of the public goods character of these resources, a

417   cultural commons must manage both use and production of cultural resources. […] This

418   characteristic of cultural commons produces a more intertwined set of exogenous variables

419   because separating the managed resources from the attributes and rules-in-use of the

420   community that produces them is impossible."

421 Genetic resources for food and agriculture lie somewhere between cultural and natural resources

422 commons (Halewood, 2013). The original raw materials were naturally occurring plants; they have since

423 been dramatically altered over the course of millennia through combinations of both natural and human

424 selection pressures. One social dilemma unique to cultural commons (and hybridised natural and

425 cultural commons such as Plant Genetic Resources for Food and Agriculture) is that underuse, not

426 overexploitation, threatens their creation and continued existence. Farmer-bred crop varieties that are

427 not maintained through continued selection will degrade and cease to exist. If they are not actively

428 reproduced, they will be subject to genetic drift and recombination with other populations until they

429 eventually cease to be what they were (see Notes S4). The same is true of most varieties bred by

430 professional plant breeders.

431 Prior to the bioinformatics boom, the range of actors involved in the generation, conservation,

432 improvement and use of plant genetic resources was already extensive. It involved farmers in centres of

433 genetic diversity; public and private sector researchers and plant breeders; and community, national

434 and international genebanks, with the Svalbard Global Seed Vault as the 'safety back-up of last resort'.

435 Halewood's (2013) observation that, "given this complexity, it is perhaps not surprising that the modular

436 organization of PGRFA commons have evolved into separate tiers, involving like-minded and like-

437 situated […] communities of actors, with various levels of connection between them" highlights the

438 pervasive tendency toward homophily within social networks. That is, the infrastructure for the

439 conservation and production of PGRFA integrates and rewards some tiers better than others. One tier is

440 organised around gene banks throughout the world. Actors within this tier are generally preoccupied

441 with and rewarded for conservation of existing genetic resources, rather than production of novel

442 genetic resources. This may change if pre-breeding activities become the remit of gene banks,

443 particularly with increased sequencing and data processing capacity. Primary responsibility for

444 innovation and production of PGRFA rests with plant breeders and farmers, both of whom have well-

445 established reward systems. Professional plant breeders tend to occupy a different tier than farmers

446 (who were the original plant breeders), one in which market forces, at least for some crops, are playing

447 an increasingly important role. There are also functional links between gene banks and breeders. The

448 situation with farmers is very different, particularly resource-poor farmers in centres of crop genetic

449 diversity. Resource-poor farmer networks are generally small and local, with poor connections to

450 national and international gene banks, and few links to professional plant breeders. Planting decisions

451 for these farmers are nonetheless often largely market driven, and/or for private consumption. Their

452 collective contributions to crop diversity over several millennia have been made without any established

453　mechanisms of reward or recognition for individual farmers or even farmers' collectives as innovators or

454　as producers of PGRFA beyond the reward of the crop harvest itself. Hodgkin et al. (2013) draw similar

455　conclusions with respect to their analysis of the state of the 'Global System for the Conservation and

456　Sustainable Use of PGRFA' developed under the auspices of the United Nations. Over the last 40 years,

457　considerable effort has been made to support *ex situ* conservation (gene banks) and professional plant

458　breeding; comparatively little has been done to support on-farm (*in situ*) conservation and local

459　breeding efforts. To date, efforts to integrate these modular architectures into cooperative networks

460　have not been successful in fostering new collaborations or greater trust among disparate interest

461　groups.

462　***Brokerage platforms to enhance innovation systems***

463　The literature on innovation theory, particularly on the role of research in agricultural innovation,

464　provides insights about the influence of organisational structure and the importance of enhanced

465　linkages between the key actors involved. The starting point for much of this literature is a rejection of

466　the idea that agricultural innovation follows a predictable, politically neutral, linear pathway controlled

467　by research centres that transfer beneficial technologies to passive recipient farmers in a 'trickle-down'

468　framework (see Notes S5). Instead, it is argued that agricultural innovation derives from multiple

469　sources, follows unpredictable paths, and is highly influenced by (if not inseparable from) institutional,

470　economic and political factors (Chambers, 1983, 2008; Biggs, 1990; Hall *et al.*, 2003, 2005). Douthwaite

471　*et al.* (2003) characterise this duality as positivism versus constructivism. Regardless of the fact that the

472　narrative of the linear innovation pathway does not accurately reflect reality, it is perpetuated because,

473　among other things, it is appealingly simple; reinforces the dominant position of scientific research

474　centres and their access to funds; responds to donors' need for predictable, measurable returns on

475　investments; and cannot be easily challenged by those outside the major research centres who are

476　actively contributing to innovation processes.

477　One focus of innovation systems research is the analysis of empirical evidence of different groups'

478　contributions to the development, diffusion, and adoption of agricultural technologies to establish the

479　extent to which those processes are centralised or decentralised. Another focus is the identification of

480　ways of increasing innovation capacity by enhancing engagement between stakeholders in the

481　development and use of technologies, with a focus on institutional learning, and strengthening the

482　linkage of previously marginalised stakeholders to the innovation process. Various studies have

483　confirmed the utility of such interventions in helping previously unorganised groups, or groups

484   disconnected from formal research scientists, to articulate their interests or demands; in lowering levels

485   of uncertainty in the preliminary stages of innovative processes; in forging alliances for new innovation

486   agendas between actors that would not or could not risk engaging in new activities on their own; and in

487   mediating conflicts among partners regarding funding allocation and intellectual property ownership

488   (Klerkx *et al.*, 2009). Examples of bodies that have enabled such progress include creation and

489   empowerment of innovation intermediaries (Howells, 2006), innovation brokers (Klerkx *et al.*, 2009),

490   multi-stakeholder platforms (Adekunle & Fatunbi, 2012), and education and training networks (Spielman

491   *et al.*, 2008). These studies have also shed light on circumstances that can lead to failure of such

492   interventions including perceptions that the broker: a) is not sufficiently independent and too closely

493   tied to the objectives of his or her own organisation, b) may eventually enter into competition with

494   other stakeholders in the innovation processes concerned, or c) has too much influence as a source of

495   expert knowledge and ends-up occupying the position of a consultant providing expert opinions,

496   undermining the group dynamic and reducing contributions from others.

497   **Figure 2.**

498   ***Strengthening network ties for innovation and policy development***

499   The 'modular architectures' described above are conceptually similar to what is known in network

500   theory as a collaborative innovation network. Networks come in many forms, but fundamentally

501   comprised of actors and relations. Actors, referred to as nodes in network theory, are differentiated by

502   influence (e.g. power, prestige) and relationships vary along an informal-formal continuum.

503   Properly functioning, participatory networks foster trust, largely through the formation of social capital.

504   Local actors, particularly farmers and community organisations, currently enjoy relatively little social

505   capital in existing PGRFA networks. Access to influential non-governmental organisations (NGOs),

506   scientific knowledge networks, and direct linkages to financial stakeholders could substantially expand

507   the social capital of local stakeholders and, importantly, enhance trust within the network.

508   Overcoming the barriers to trust, and the divergence of interest, between the different actors interested

509   in PGFRA is central for accelerating the development of improved crops. However, it will require new

510   models of scientific practice that redefine the traditional top-down (hierarchical) models that have

511   dominated the field. Greater engagement with stakeholders who have traditionally been implicitly or

512   explicitly viewed as passive, peripheral participants in the larger innovation process is critical if we are to

513   break out of the current situation where material-sharing is in decline and benefit-sharing is not

514   apparent.

515   We briefly mention two types of networks that could inform the structure and *modus operandi* of new

516   genetic resources or bioinformatics-based innovation platforms. Peer network academies (Klerkx &

517   Leeuwis, 2009) represent a high-value collaborative network hub due to their focus on networks of

518   stakeholders that share common interests (e.g. corn, rice, dairy, and poultry in the agricultural industry).

519   Two key features of a peer network academy organised around PGRFA are an online databank, which

520   can serve as a clearinghouse of essential information, including germplasm input systems linking to

521   larger data repositories, and a collaborative structure in which farmers are active participants in

522   knowledge discovery by virtue of their fields functioning as demonstration farms (citizen scientists).

523   A second type of network model to be considered in this context is that of Collaborative Innovation

524   Networks (CoINs). CoINs facilitate inter-network collaborations that span traditional social, economic,

525   and cultural hierarchies and boundaries to encourage direct communication between actors that have

526   traditionally had little or no direct communication and collaboration (Gloor, 2006). They are

527   characterised by a widely dispersed but interdependent membership working toward common goals in

528   an environment of trust. A hallmark of CoINs is their lack of central management, which allows broadly-

529   based transparent interactions among network actors. However, new innovation platforms will likely (at

530   least initially) require investment in more centralised governance mechanisms to identify and promote

531   shared goals and trust among the range of actors that we underscore need to be involved. Of course,

532   some of the 'modules' of the innovation platform could embrace CoINs-inspired structures and modus

533   operandi, and over time the proportion of such activities related to the innovation platform could

534   increase.

535   To address hierarchical organisational structures and concentrated power among elite actors involved in

536   genetic resources or bioinformatics-based innovation platforms, we suggest two possible approaches.

537   Farmers, especially those in developing countries, will need a seat at the table that is not merely

538   symbolic, but functional. Redefining farmers as citizen-scientists filling an integral role in field

539   experimentation and data generation as part of a modular, commons-based innovation system holds

540   tremendous potential to overcome the historical backdrop of mistrust between local stakeholders and

541   elites working in agri-business, science, and public policy. Also, data generating systems could and

542   should link field data, germplasm information, and relevant metadata in a manner that moves farmers

543   from the end of the conventional agricultural research extension pipeline (Klerkx *et al.*, 2009) to a

18

544   central, co-equal role in the PGR collaborative innovation network. Farmers (and the system as a whole)

545   should benefit from being able to develop new, collaborative linkages with scientists, civil society

546   organisations, and agri-industry groups. Such a data network would require feedback loops that

547   facilitate not only the transmission of information into large-scale data repositories (e.g. genetic

548   sequence, phenomic, environmental, and GIS data) but also cycle information back out to the field in an

549   open and transparent manner that engenders trust, furthers cooperation, and produces equitable

550   benefits across the network. Complementary capacity strengthening is necessary to promote the ability

551   of farmers organisations and resource-poor national agricultural research and extension services to take

552   advantage of these networks.

553   ## 5. Access and benefit sharing and big data

554   As highlighted above, much of the support for international access and benefit sharing (ABS) laws came

555   from developing countries who were (and are) concerned about existing inequities in the distribution of

556   benefits derived from commercial use of genetic resources. If institutionalised, the modular, inclusive,

557   governance mechanisms described above could help allay concerns that the new bioinformatics

558   capacities will exacerbate those inequities. To increase the likelihood of this positive outcome,

559   organisations seeking to catalyse new genetic resources/bioinformatics-based innovation platforms

560   should promote best practices and develop voluntary standards explicitly addressing ABS issues.

561   Demonstrable compliance with these standards could be a precondition of endorsement by innovation

562   platform(s) of project proposals prepared for donors, or projects in which platforms engage directly. The

563   advantage of this approach to developing ABS standards and best practices is that they can be

564   developed organically, building on existing practices and reciprocal interests of the actors involved. In

565   this way, they could be sufficiently flexible to take into account the very different motivations of the

566   very different groups of actors, and the wide range of both non-monetary and monetary benefits that

567   can be generated.

568   Meanwhile, as discussed above, the issue of sharing benefits derived from the use of genetic sequence

569   data has already made its way onto a number of international agendas, with developing regions calling

570   for new, internationally negotiated benefit-sharing rules. The ITPGRFA Governing Body, Conference of

571   the Parties to the CBD, and FAO (Food and Agriculture Organisation) Commission on Genetic Resources

572   for Food and Agriculture have all initiated fact-finding processes for considering the impact of genome

573   sequencing and synthesis on the conservation and sustainable use of genetic resources, and the

574   equitable sharing of benefits. The issue is also being considered under the framework of the WHO

575   (World Health Organisation) Pandemic Influenza Preparedness (PIP) framework. The PIP Framework

576   Advisory Group and a specialized technical working group are considering options for benefit sharing

577   linked to the sharing and commercial use of gene sequence data of influenza viruses with human

578   pandemic potential (see Notes S2). On one hand, there is clearly a need for novel approaches to

579   promote both monetary and non-monetary benefit sharing. On the other hand, we are concerned that

580   discussions at the level of the United Nations have historically tended to focus almost exclusively on

581   monetary benefit-sharing (largely overlooking ways to promote valuable forms of non-monetary

582   benefit-sharing).

583   As the summary of the last 30 years of international policy making above underscores, it is challenging

584   to develop globally applicable, legally binding ABS norms that are custom fit for emerging areas of

585   scientific practice. There is a risk that new efforts at the level of the United Nations to develop a one-

586   size-fits-all, ABS policy solution with respect to genomic sequence data could inadvertently end-up

587   perpetuating disincentives for sharing, accessing, and using genetic resources and information (including

588   genomic sequence data). New rules, if inappropriately crafted, could inadvertently create barriers to the

589   development of innovation platforms and enhanced governance arrangements as described above. In

590   the following paragraphs, we briefly consider a range of options regarding new ABS policies that are,

591   have been, or could be, considered by these intergovernmental bodies. This is not meant to constitute a

592   thorough analysis of ongoing negotiations; only to provide an introductory insight into the kinds of

593   benefit sharing 'solutions' that are actively under consideration in those fora.

594   Perhaps the lightest-weight option would be for one, or some combination, of those international

595   bodies to opt for a 'soft' norm approach, without the creation of new, legally binding obligations. This

596   approach could involve identifying and endorsing best practices, developing voluntary guidelines and

597   model ABS agreements. It could also include self-reporting mechanisms through which new innovation

598   platforms could report on their management structures, guiding principles, partnerships and activities,

599   and seek endorsements from relevant international bodies.

600   Another approach would entail extending the scope of the Nagoya Protocol and the ITPGRFA to apply to

601   genome sequence and other types of digital data related to PGR, in addition to material genetic

602   resources. Under the ITPGRFA, this would entail data owners and curators agreeing to provide

603   facilitated access to data sets subject to the condition that if that information was 'incorporated' (or

604   used) in the development of new, commercialised PGRFA products, the data user would have to make

605 payments to the Benefit-Sharing Fund (BSF). Under the predominant model for national implementation

606 of Nagoya, parties seeking access to a database containing genome sequence or other PGR-related

607 digital data would need to negotiate an ABS agreement, which could include any number of conditions.

608 The same mechanisms for monitoring and enforcing compliance with ABS agreements for material

609 genetic resources would be extended to trace and enforce agreements related to digital PGR-associated

610 data. Unfortunately, as highlighted above, these contract-based, track-and-trace ABS systems are

611 already proving to be very difficult to implement given the non-rivalrous and non-excludable nature of

612 material genetic resources. Controlling access to information – given that it is already so diffused, easily

613 copied, and easily moved through the Internet – would be even more difficult, as would be tracking and

614 tracing the use of that information in the development of a discrete, new, commercialised crop variety

615 or patented trait. One can imagine the adoption of such a model eventually leading to collapse of the

616 ABS systems that the international community has been working to develop since the CBD came into

617 force. On the other hand, perhaps the technological breakthroughs that have created enhanced interest

618 and value for genetic sequence information could also be deployed in enhanced systems for monitoring

619 and verifying uses of those data.

620 A more pragmatic variation of the contract approach that is currently being considered under the

621 ITPGRFA framework is to create a subscription system for accessing both genetic resources and

622 sequence data, and for sharing related monetary benefits. Under this system, commercial users (or

623 governments representing clusters of commercial users) would commit to making annual payments to

624 the BSF, based on their annual seed sales, for a fixed period (e.g. ten years). During that time, they

625 would have facilitated access to both genetic resources in the multilateral system and also genetic

626 sequence data. Since the rate of payment is based on seed sales generally, there would be no need to

627 track and trace the use of the material genetic resources or sequence data in the creation of new

628 products. This model is attractive in that it could generate more predictable levels of funding. It would

629 also simultaneously address ABS commitments for both genetic resources and digital sequence data,

630 and in a way that significantly reduces transaction costs associated with the 'pure' contractual model

631 currently in place for materials.

632 From the point of view of non-governmental actors involved in conserving and using genetic resources

633 and genomic data, the most straightforward option would be for national governments to undertake to

634 make financial contributions to the BSF on a percentage of seed sales within their borders, without

635 linking the actual incorporation of the genetic resource or the data in new commercialised products. In

636   return, natural and legal persons within their borders could enjoy facilitated access to both crop genetic
637   resources in the multilateral system of ABS and digital genetic sequence data. The system would also
638   not require tracking or tracing of the use of materials or information. National governments could
639   decide whether or not to recoup those costs from their own seed companies. There is a precedent for
640   this model; Norway has adopted the policy of voluntarily making payments to the BSF based on 0.1% of
641   seed sales. In the context of the on-going renegotiations of the ABS conditions of the ITPGRFA's
642   multilateral system, some developing countries and observer organisations have promoted this
643   approach. A number of developed countries have rejected it. Ultimately, in order to avoid the
644   extraordinary complexities that would be associated with a legally binding extension of the contractual
645   model to genome sequence data, it is possible that the practical merits of this approach will be more
646   widely appreciated.

647   Under the Nagoya Protocol, some have argued that it would also be possible to develop similar
648   multilateral ABS arrangements for some classes of genetic resources and related information, though to
649   date, there have not been any concrete initiatives to do so (see Notes S7).

650   **6. Conclusion**

651   It will most certainly take several years for the international community to develop mechanisms to
652   address the issues raised in this review, particularly if it is collectively decided that new legally binding
653   agreements (or amendments or protocols to existing legally binding agreements) are necessary. In the
654   meantime, there will be opportunities for interested organisations and networks to develop inclusive
655   forms of governance for the deployment of the new technical capacities discussed in this paper to
656   realise the sustainable development goals. To succeed, broad coalitions of scientists, information
657   technologists, gene bank managers, breeders, farmers and civil society organisations will need to find
658   opportunities to articulate a set of common goals and develop inclusive, transparent, systems for
659   working together. If they are successful, the governance mechanisms, best practices and benefit-sharing
660   standards they develop could positively influence the tone of on-gong intergovernmental negotiations
661   and the form and content of norms that are eventually developed under the aegis of the United Nations.
662   The ball is now in the court of champions of these new technologies to foster innovation platforms and
663   governance systems, which will inspire trust and promote the most effective, equitable deployment of
664   those technologies.

674   **ORCID**

675   Wayne Powell: http://orcid.org/0000-0002-5612-3398

676   **References**

677   **Adekunle AA, Fatunbi AO**. **2012**. Approaches for setting-up multi-stakeholder platforms for agricultural

678   research and development. *World Applied Sciences Journal* **16**: 981–988.

679   **Authors: Toronto International Data Release Workshop**. **2009**. Prepublication data sharing. *Nature* **461**:

680   168–170.

681   **Biggs S**. **1990**. A Multiple Source of Innovation Model of Agricultural Research and Technology

682   Promotion. *World Development* **18**: 1481–1499.

683   **Cavanagh C, Morell M, Mackay I, Powell W**. **2008**. From mutations to MAGIC: resources for gene

684   discovery, validation and delivery in crop plants. *Current Opinion in Plant Biology* **11**: 215–221.

685   **Chambers R**. **1983**. *Rural Development: Putting the Last First*. London: Longman.

686   **Chambers R**. **2008**. *Revolutions in Development Inquiry*. London: Earthscan.

687   **Collard BCY, Mackill DJ**. **2008**. Marker-assisted selection: an approach for precision plant breeding in the

688   twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**: 557 LP-

689   572.

690   **Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, Wright D, Graham G**. **2014**.

691   Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction.

692   *Crop and Pasture Science* **65**: 311–336.

693   **Dedeurwaerdere T**. **2013**. Institutionalizing global genetic resource commons for food and agriculture.

694   In: Halewood M, Lopez Noriega I, Louafi S, eds. Crop Genetic Resources as a Global commons.

695   Challenges in International Law and Governance. Oxford: Routledge, 369–391.

696   **Douthwaite B, Kuby T, van de Fliert E, Schulz S**. **2003**. Impact pathway evaluation: an approach for

697   achieving and attributing impact in complex systems. *Agricultural Systems* **78**: 243–265.

698   **Ersoz ES, Yu J, Buckler ES**. **2009**. Applications of Linkage Disequilibrium and Association Mapping in

699   Maize. In: Kriz AL, Larkins BA, eds. Molecular Genetic Approaches to Maize Improvement. Berlin,

700   Heidelberg: Springer Berlin Heidelberg, 173–195.

701   **van Etten J, Beza E, Calderer L, Van Duijvenkijk K, Fadda C, Fantahun B, Kidane YG, Van De Gevel J,**

702   **Gupta A, Mengistu DJK,** *et al.* **2016**. First experiences with a novel farmer citizen science approach:

703   Crowdsourcing participatory variety selection through on-farm Triadic Comparisons of Technologies

704   (TRICOT). *Experimental Agriculture* **First View**: 1–22.

705   **van Etten J, Steinke J, van Wijk M**. **2017**. How can the Data Revolution contribute to climate action in

706   smallholder agriculture. *Agriculture for Development* **30**: 44–48.

707   **Furbank RT, Tester M**. **2011**. Phenomics-technologies to relieve the phenotyping bottleneck. *Trends in*

708   *plant science* **16**: 635–644.

709   **Gloor PA**. **2006**. *Swarm Creativity: Competitive Advantage through Collaborative Innovation Networks*.

710   Oxford: Oxford University Press.

711   **Goodwin S, McPherson JD, McCombie WR**. **2016**. Coming of age: ten years of next-generation

712   sequencing technologies. *Nature Review Genetics* **17**: 333–351.

713   **Halewood M**. **2013**. What kind of goods are plant genetic resources for food and agriculture? towards

714   the identification and development of a new global commons. *International Journal of the Commons* **7**:

715   278–312.

716   **Hall A, Mytelka L, Oyeyinka B**. **2005**. *Innovation systems: Implications for agricultural policy and*

717   *practice. Institutional Learning and Change (ILAC) Brief—Issue 2* (null, Ed.). Rome: Bioversity

718   International.

719   **Hall A, Rasheed Sulaiman V, Clark N, Yoganand B**. **2003**. From measuring impact to learning

720   institutional lessons: an innovation systems perspective on improving the management of international

721   agricultural research. *Agricultural Systems* **78**: 213–241.

722   **Hamblin MT, Buckler ES, Jannink J-L**. **2011**. Population genetics of genomics-based crop improvement

723   methods. *Trends in Genetics* **27**: 98–106.

724   **Hammond E**. **2016**. *Digital genebankers plan to ignore UN request on the impact of genomics and*

725   *synthetic biology on access and benefit sharing: A preliminary report*. Penang: Third World Network.

726   **Hodgkin T, Demers N, Frison E**. **2013**. The evolving global system of conservation and use of plant

727   genetic resources for food and agriculture: what is it, and where does the Treaty fit it? In: Halewood M,

728   Lopez Noriega I, Louafi S, eds. Crop Genetic Resources as a Global Commons. Oxford: Routledge, 344–

729   367.

730   **Howells J**. **2006**. Intermediation and the role of intermediaries in innovation. *Research Policy* **35**: 715–

731   728.

732   **James C**. **2014**. *Global Status of Commercialized Biotech/GM Crops: 2014*. Ithaca, NY: International

733   Service for the Acquisition of Agri-biotech Applications (ISAAA).

734   **Klerkx L, Hall A, Leeuwis C**. **2009**. Strengthening agricultural innovation capacity: are innovation brokers

735   the answer? *International Journal of Agricultural Resources, Governance and Ecology* **8**: 409–438.

736   **Klerkx L, Leeuwis C**. **2009**. The emergence and embedding of innovation brokers at different innovation

737   system levels: Insights from the Dutch agricultural sector. *Technological Forecasting and Social Change*

738   **76**: 849–860.

739   **Lawson C, Rourke M**. **2016**. Open Access DNA , RNA and Amino Acid Sequences : The Consequences and

740   Solutions for the International Regulation of Access and Benefit Sharing. *Journal of Law & Medicine* **24**:

741   96–118.

742   **Lipka AE, Kandianis CB, Hudson ME, Yu J, Drnevich J, Bradbury PJ, Gore MA**. **2015**. From association to

743   prediction: statistical methods for the dissection and selection of complex traits in plants. *Current*

744   *Opinion in Plant Biology* **24**: 110–118.

745 **Madison MJ, Frischmanrrff BM, Strandburg KJ**. **2010**. Constructing commons in the cultural

746 environment. *Cornell Law Review* **95**: 657–709.

747 **McCouch SR, McNally KL, Wang W, Sackville Hamilton R**. **2012**. Genomics of gene banks: A case study

748 in rice. *American Journal of Botany* **99**: 407–423.

749 **Moore GE**. **1965**. Cramming More Components onto Integrated Circuits. *Electronics* **38**: 114–117.

750 **Noman A, Aqeel M, He S**. **2016**. CRISPR-Cas9: Tool for Qualitative and Quantitative Plant Genome

751 Editing. *Frontiers in Plant Science* **7**: 1740.

752 **Ostrom E, Hess C**. **2010**. *Private and Common Property Rights*. Cheltenham: Edward Elgar Publishing.

753 **Pandemic Influenza Preparedness Framework Advisory Group's Technical Working Group on the**

754 **Sharing of Influenza Genetic Sequence Data**. **2016**. *Optimal Characteristics of an influenza genetic*

755 *sequence data sharing system under the PIP Framework*. Geneva: WHO - Pandemic Influenza

756 Preparedness (PIP) Working Group.

757 **Petolino JF, Srivastava V, Daniell H**. **2016**. Editing Plant Genomes: a new era of crop improvement. *Plant*

758 *Biotechnology Journal* **14**: 435–436.

759 **Poland J**. **2015**. Breeding-assisted genomics. *Current Opinion in Plant Biology* **24**: 119–124.

760 **Policy Science for Environment**. **2016**. *Synthetic biology and biodiversity. Future Brief 15.* Produced for

761 the European Commission DG Environment by the Science Communication Unit, UWE, Bristol.

762 **Puchta H**. **2017**. Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Current*

763 *Opinion in Plant Biology* **36**: 1–8.

764 **Reichman JH, Uhlir PF, Dedeurwaerdere T**. **2016**. *Governing Digitally Integrated Genetic Resources,*

765 *Data, and Literature: Global Intellectual Property Strategies for a Redesigned Microbial Research*

766 *Commons*. Cambridge: Cambridge University Press.

767 **Sardos J, Rouard M, Hueber Y, Cenci A, Hyma KE, van den Houwe I, Hribova E, Courtois B, Roux N**.

768 **2016**. A Genome-Wide Association Study on the Seedless Phenotype in Banana (Musa spp.) Reveals the

769 Potential of a Selected Panel to Detect Candidate Genes in a Vegetatively Propagated Crop. *PLOS ONE*

770 **11**: e0154448.

771   **Satizábal H, Barreto-Sanz M, Jiménez D, Pérez-Uribe A, Cock J**. **2012**. Enhancing Decision-Making

772   Processes of Small Farmers in Tropical Crops by Means of Machine Learning Models BT - Technologies

773   and Innovations for Development: Scientific Cooperation for a Sustainable Future. In: Bolay J-C, Schmid

774   M, Tejada G, Hazboun E, eds. Paris: Springer Paris, 265–277.

775   **Schiml S, Puchta H**. **2016**. Revolutionizing plant biology: multiple ways of genome engineering by

776   CRISPR/Cas. *Plant Methods* **12**: 8.

777   **Spielman DJ, Ekboir J, Davis K, Ochieng CMO**. **2008**. An innovation systems perspective on

778   strengthening agricultural education and training in sub-Saharan Africa. *Agricultural systems* **98**: 1–9.

779   **Stevenson JR, Villoria N, Byerlee D, Kelley T, Maredia M**. **2013**. Green Revolution research saved an

780   estimated 18 to 27 million hectares from being brought into agricultural production. *Proceedings of the*

781   *National Academy of Sciences* **110**: 8363–8368.

782   **The International Civil Society Working Group on Synthetic Biology**. **2016**. *Synthetic Biology and the*

783   *CBD: Five key decisions for COP 13 & COP-MOP 8*. [WWW document] URL

784   https://www.boell.de/sites/default/files/2016-11-icswgsb_synbio_brief_cop13_.pdf. [accessed 13

785   January 2018].

786   **Walker JM, Gardner R, Ostrom E**. **1990**. Rent dissipation in a limited-access common-pool resource:

787   Experimental evidence. *Journal of Environmental Economics and Management* **19**: 203–211.

788   **Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W,**

789   **da Silva Santos LB, Bourne PE, *et al.* 2016**. The FAIR Guiding Principles for scientific data management

790   and stewardship. **3**: 160018.

791   **Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P, Hu L, Yamasaki M, Yoshida S, Kitano H, Hirano K, *et al.***

792   **2016**. Genome-wide association study using whole-genome sequencing rapidly identifies new genes

793   influencing agronomic traits in rice. *Nat Genet* **48**: 927–934.

794   **Figure Legends**

795   **Figure 1. Schematic presentation of a model for discovery and breeding platforms.** Multi-environment

796   testing and information or knowledge sharing underpin systematic assessment and improvement of

797   plant genetic resources to produce elite germplasm and new products

798

799 **Figure 2. Generalised governance structure.** ⟶ = knowledge-driven forward flow of materials,

800 technology and information to benefit the farmer. ⟵ = reverse-flow of data and information to

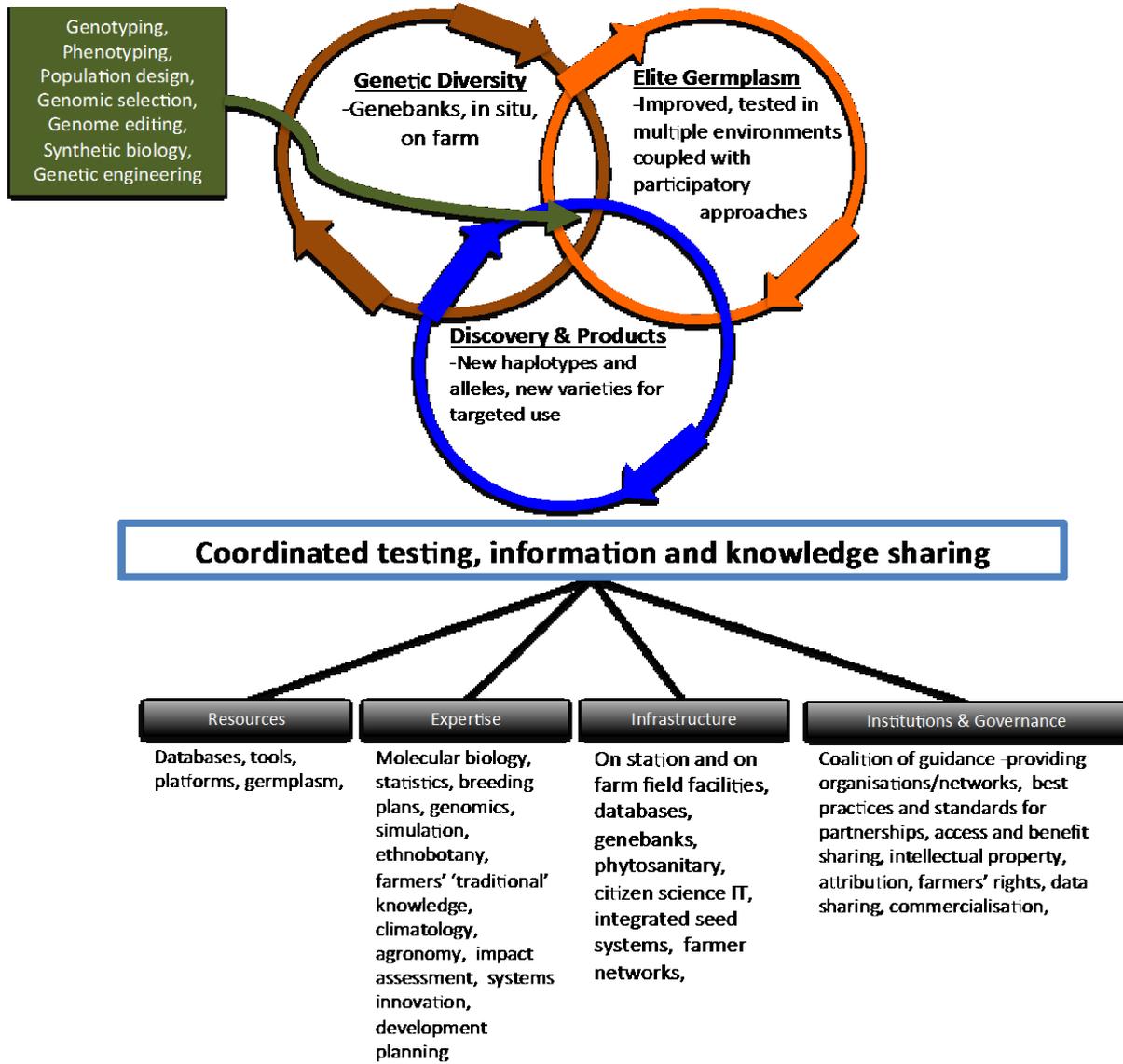801 enhance earlier steps in the impact pathway.

802

Fig. 1

Genotyping, Phenotyping, Population design, Genomic selection, Genome editing, Synthetic biology, Genetic engineering

Genetic Diversity
-Genebanks, in situ, on farm

Elite Germplasm
-Improved, tested in multiple environments coupled with participatory approaches

Discovery & Products
-New haplotypes and alleles, new varieties for targeted use

Coordinated testing, information and knowledge sharing

Resources	Expertise	Infrastructure	Institutions & Governance
Databases, tools, platforms, germplasm,	Molecular biology, statistics, breeding plans, genomics, simulation, ethnobotany, farmers' 'traditional' knowledge, climatology, agronomy, impact assessment, systems innovation, development planning	On station and on farm field facilities, databases, genebanks, phytosanitary, citizen science IT, integrated seed systems, farmer networks,	Coalition of guidance -providing organisations/networks, best practices and standards for partnerships, access and benefit sharing, intellectual property, attribution, farmers' rights, data sharing, commercialisation,

Fig. 2

Gaps & opportunities assessments ,
policy guidance to advance SDGs:
coalition of donors,
public and private R&D orgs,
farmer networks, reps from
developed and developing countries

International and national policies, programmes, priorities

Conservation

On farm use

R&D partnerships:
public / private / farmer;
international /
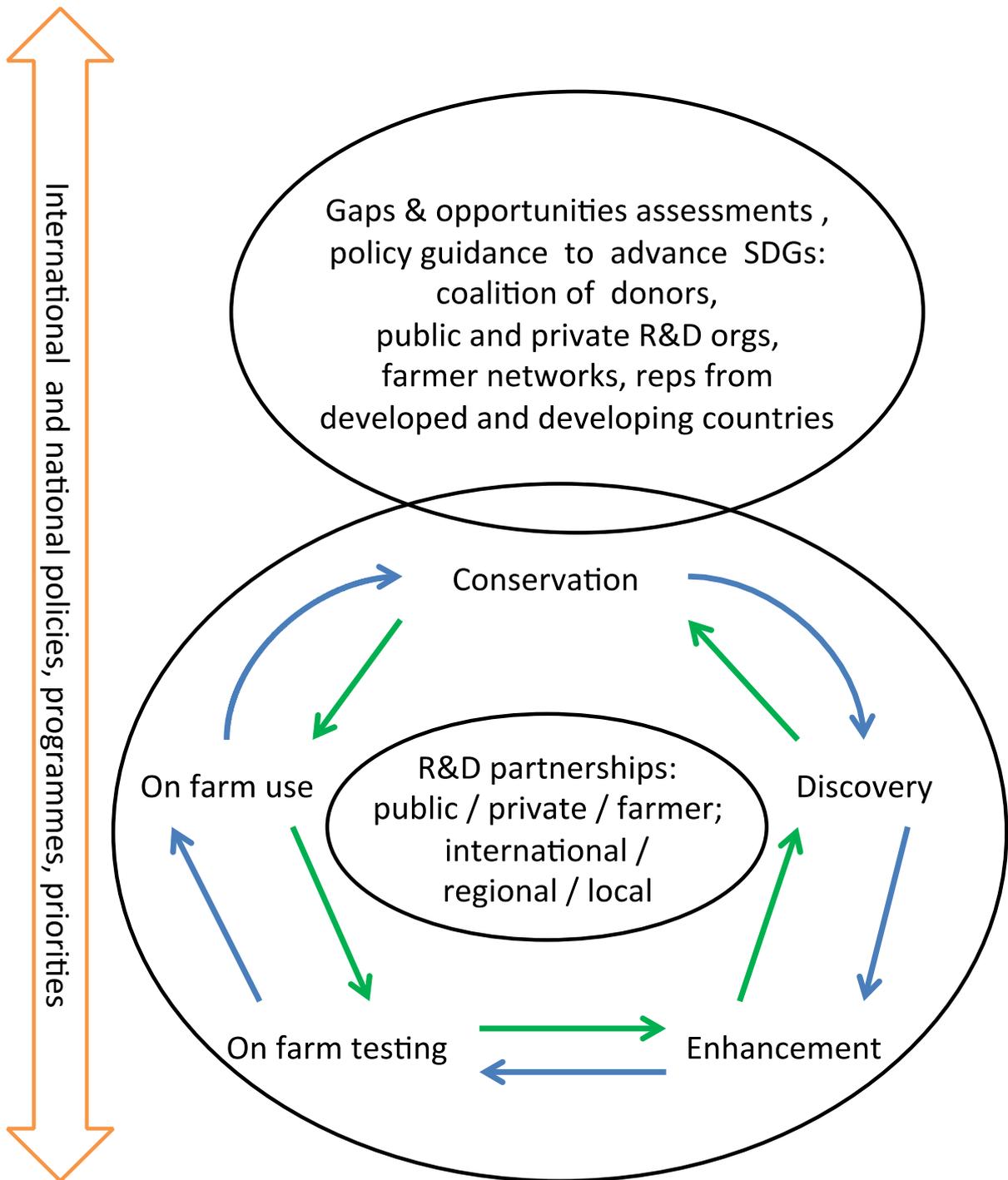regional / local

Discovery

On farm testing

Enhancement

New Phytologist **Supporting Information**

Article title: *Plant genetic resources for food and agriculture: opportunities and challenges emerging from the science and information technology revolution*

Authors: *Michael Halewood, Tinashe Chiurugwi, Ruaraidh Sackville Hamilton, Brad Kurtz, Emily Marden, Eric Welch, Frank Michiels, Javad Mozafari, Muhamad Sabran, Nicola Patron, Paul Kersey, Ruth Bastow, Shawn Dorius, Sonia Dias, Susan McCouch, and Wayne Powell*

Article acceptance date: 21 November 2017

The following Supporting Information is available for this article:

Description/Legends for supporting information notes

Notes S1 Expansion of the discussion on potential IP claims with respect to data made available on GenBank.

Notes S2 Further details about previous discussions concerning potential inclusion of information about PGR, e.g. genomic sequence or phenotypic data, in the Nagoya protocol.

Notes S3 Introduction to an example (from a WHO initiative) of an attempt to address concerns about benefit sharing obligations following commercial use of open-access genetic sequence data.

Notes S4 Some farmers in crop centres of diversity are seeking recognition for their roles in both the creation and further exploitation of PGR and related data.

Notes S5 Further details about an alternative view on the creation and dissemination of agricultural innovation.

Notes S6 An alternative view on how multilateral access and benefit sharing arrangements for PGR and related information could be achieved under the Nagoya Protocol.

Notes S1 Lawson and Rourke (2016) note that there has been a gradual evolution in the treatment of potential IP claims with respect to data made available on GenBank, EBML-DBI and DDBJ with some erosion of the original open access principles embraced by those repositories. Among other evidence, they reproduce the NCBI/Genbank policy that it "cannot provide comment or unrestricted permission concerning the use, copying or distribution of the [data and] information contained in Genebank" and that anyone "reproducing, redistribution, or making commercial use of this information are expected to adhere to the terms and conditions asserted by the copyright holder."

Notes S2 At one point during the negotiations of the Nagoya Protocol it was proposed by some delegates that 'information or knowledge of value derived from genetic material ' should be included in the definition of 'derivatives' to be regulated by the Nagoya Protocol (Secretariat of the Convention on Biological Diversity, 2008). However, that proposal was not accepted, and information was not included in the final text. Nonetheless, it has been argued by some that the Nagoya Protocol can be understood as extending in scope to genomic sequence data through a broad interpretation of the term 'utilisation' in the Nagoya Protocol ("to conduct research and development on the genetic and or biological composition of genetic resources") For expanded discussion of this issue, see Manzella (2016).

Notes S3 The consideration of this issue has a longer history under the framework of the WHO Pandemic Influenza Preparedness (PIP) framework adopted in 2011. Section 5.2.4 of the PIP Framework states that the Director General of WHO should consult with the PIP Advisory Group

concerning resolution of issues related to the handling of the genetic sequence data (GSD) of influenza viruses with pandemic potential. Since 2013, the Advisory Group and later a specialised Technical Working Group have been considering options for benefit sharing linked to the sharing and commercial use of gene sequence data of influenza viruses with human pandemic potential. It has generated background papers concerning practical approaches for establishing and monitoring benefit sharing standards, see, for example, PIP Advisory Group Technical Working Group on the Sharing of Influenza Genetic Sequence Data (2016).

Notes S4 And one of the social dilemmas facing the use of genetic resources sequence data is that the 'rules-in-use' of at least part of the community that played a role in producing them (i.e. farmers in centres of crop diversity) appear to require more active recognition of their contributions and engagement in systems for their further exploitation.

Notes S5 Sumberg et al. (2003) argue that Biggs exaggerates aspects of what he calls the 'central source model' overlooking, for example, that agricultural extension was intentionally created with the idea of facilitating contributions to innovation from extension agents and others involved in the diffusion and adoption of technologies.

Notes S6 Nagoya Protocol article 10 states: "Parties shall consider the need for and modalities of a global multilateral benefit-sharing mechanism to address the fair and equitable sharing of benefits derived from the utilization of genetic resources and traditional knowledge associated with genetic resources that occur in transboundary situations or for which it is not possible to grant or obtain prior informed consent. The benefits shared by users of genetic resources and traditional knowledge associated with genetic resources through this mechanism shall be used to support the conservation of biological diversity and the sustainable use of its components

globally". It has been suggested (C. Lawson, pers comm) that a multilateral system for genomic sequence data could be modelled on existing copyright collection societies, with payment rates based on uses of sequence data that are estimated by way of on-line searches or audits of final products.

References

Lawson C, Rourke M. **2016**. Open Access DNA , RNA and Amino Acid Sequences : The Consequences and Solutions for the International Regulation of Access and Benefit Sharing. *Journal of Law & Medicine* **24**: 96–118.

Manzella D. **2016**. *The Global Information System and Genomic Information: Transparency of Rights and Obligations: Background Study Paper 10 (IT/GB7/SAC-1/16/ BSP 10) prepared for the First meeting of the Scientific Advisory Committee on Article 17 of the International Treat*. Rome: FAO International Treaty on Plant Genetic Resources for Food and Agriculture.

Pandemic Influenza Preparedness Framework Advisory Group's Technical Working Group on the Sharing of Influenza Genetic Sequence Data. **2016**. *Optimal Characteristics of an influenza genetic sequence data sharing system under the PIP Framework*. Geneva: WHO - Pandemic Influenza Preparedness (PIP) Working Group.

Secretariat of the Convention on Biological Diversity. **2008**. *Report of the Meeting of the Group of Legal and Technical Experts on Concepts, Terms, Working Definitions and Sectoral Approaches. UNEP/CBD/ABSWG/7/2*. Montreal: Convention on Biological Diversity.

Sumberg J, Okali C, Reece D. **2003**. Agricultural research in the face of diversity, local knowledge and the participation imperative: theoretical considerations. *Agricultural systems*

76: 739–753.