



Advisory
Services



Synthesis of Learning from a Decade of CGIAR Research Programs

M. Holderness, J. Howard, I. Jouini, D. Templeton, C. Iglesias, D. Molden and N. Maxted

June 2021

Correct citation: CAS Secretariat (CGIAR Advisory Services Shared Secretariat). (2021). Synthesis of Learning from a Decade of CGIAR Research Programs. Rome: CAS Secretariat Evaluation Function.

<https://cas.cgiar.org/>

Cover image: Landscape. Credit: Elie Ntirengaya

Design and layout: Krzysztof Kaczmarek and Abhilasha Vaid

Synthesis of Learning from a Decade of CGIAR Research Programs

*M. Holderness, J. Howard, I. Jouini,
D. Templeton, C. Iglesias, D. Molden
and N. Maxted*

June 2021

Acknowledgments

This evaluation synthesis report was prepared by a team led by Mark Holderness and co-authored by Julie Howard; the evaluation techniques were implemented under the authority of evaluation expert Ibtissem Jouini. Subject matter expertise was provided by Carlos Iglesias, Deborah Templeton, David Molden, Nigel Maxted and Julie Howard (see Annex 7 in a separate document for their bios).

The synthesis was implemented under the overall direction of Allison Grove Smith, Director of the CGIAR Advisory Services (CAS). The synthesis team worked under the technical guidance and management of Svetlana Negroustoueva, CAS Evaluation Function Lead and with the support of Karen-Erdoó Jay-Yina, Senior Evaluation Officer. Analytical and content support was provided by Max Runzel and Bayardo Benítez.

The content of the report benefited from several rounds of peer review by Ravinder Kumar, John Porter, Christian Roth, Jennifer Thomson, Paolo Sarfatti, Bhavani Shankar, and Ann Waters-Bayer.

The synthesis team gratefully acknowledges the work of over 150 authors of evaluations, reviews, and other source documents upon which the synthesis is based (Annex 4). Regina Birner, lead author of the 2016 synthesis of CRP evaluations, also provided helpful insights. The synthesis team also thanks the Independent Science for Development Council (ISDC) members and CRP and platform leaders who gave their time for validating results of this synthesis and provided useful information and insights.

We wish to thank CAS staff in Rome, who provided great support to the synthesis team, ensuring the smooth conduct of technical and administrative processes related to the evaluative exercise.

Content

Executive Summary	I
1. Introduction and Methodology	I
2. Trends and Lessons across the Two Phases of CRPs.....	I
2.1 In general, CRPs have:	I
2.2 Quality of Science.....	II
2.3 Quality of Research for Development	II
2.4 Delivery of Outcomes and Sustainability.....	II
2.5 Management and Governance	IV
3. Conclusions: System-wide Issues That Have Strengthened or Weakened Achievement of CRP and CGIAR System Results	IV
3.1 CRP Funding and Its Implications	IV
3.2 Reality Check: Time, Resources, Skills, and Partnerships Required for Scaling Development Outcomes.....	V
3.3 Inadequate Monitoring and Learning for Development Outcomes	VII
4. Actionable Recommendations	VII
4.1 Recommendations for One CGIAR	VII
4.2 Recommendations for the Three Strategic Action Areas (AA).....	VIII
4.2.1 Resilient Agri-Food Systems (AA2)	VIII
4.2.2 Genetic Innovation (AA3).....	VIII
4.3 Seven Ways of Working	IX
5. Gaps and Needs for Future Evaluations.....	XII
5.1 Thematic Gaps.....	XII
5.1 Methodological Gaps.....	XII
1 Background and Synthesis Context.....	1
1.1 CGIAR Reform and Evolution of the CRPs	1
1.2 Objectives and Questions	5
1.3 Overall Approach.....	5
1.4 Limitations	7
2 Key Findings: Trends and Lessons across the Two Phases of CRPs	7
2.1 Quality of Science, Quality of Research for Development (QoR4D), Inputs, Processes, and Delivery of Outputs	7
2.1.1 Research Inputs	7
2.1.2 Research Processes and Delivery of Research Outputs	9
2.1.3 Quality of Research for Development.....	10
2.2 Performance: Delivery to Outcomes, Objectives, and Sustainability	12
2.2.1 Delivery of Outputs and Outcomes	12
2.2.2 Limited Coherence in the Use of ToC	13
2.2.3 Insufficient Research, Measures, and Skills for Sustainability and Scaling	14
2.2.4 Management and Governance	17
2.2.5 Conclusions from the Analysis of Key Findings, and Trends.....	17

3 CGIAR System-Wide Lessons and Conclusions	18
3.1 Impact Pathways of CGIAR Research	18
3.2 Monitoring, Evaluation, Learning, and Impacts Assessments (MELIA)	19
3.3 Greater Social Science Engagement to Bring Clearer Focus on the Poor and on Scaling Processes	20
3.4 Ownership of the Agenda by External Partners and Stakeholders	21
3.5 Resourcing and Measuring Contributions of the CGIAR and Partners to Development	21
4 Actionable Recommendations	21
4.1 Recommendations for One CGIAR	21
4.2 Recommendations for the Three Strategic Action Areas	22
4.2.1 Systems Transformation (AA1)	22
4.2.2 Resilient Agri-Food Systems	22
4.2.3 Genetic Innovation	23
4.3 Recommendations for Seven Ways of Working	23
4.3.1 Embracing a Systems Transformation Approach, Seeking Multiple Benefits across Impact Areas	23
4.3.2 Leveraging Ambitious Partnerships for Change	24
4.3.3 Positioning Regions, Countries, and Landscapes as Central Dimensions	24
4.3.4 Generating Scientific Evidence on Multiple Transformation Pathways	25
4.3.5 Targeting Risk Management and Resilience as Critical Qualities	25
4.3.6 Harnessing Innovative Finance	25
4.3.7 Making the Digital Revolution Central to Our Way of Working	25
5 Key Evidence Gaps	25
5.1 Thematic Gaps	26
5.2 Methodological Gaps	27

List of Figures

Figure 1. Evolution of the CRPs over two phases of implementation	3
Figure 2. Mapping of system-level outcomes and cross-cutting themes by SRFs	4
Figure 3. Five clusters for component analysis, by subject matter experts	6
Figure 4. CGIAR revenue, 2011–19 (US\$ billion)	8

List of Boxes

Box 1. What is evaluative evidence related to system-level targets?	III
Box 2. What is the evaluative evidence about CRP funding challenges?	VI
Box 3. How do Recommendations Map to Impact Areas	XI
Box 4. Bibliometric analysis	11
Box 5. Scaling in the CRPs	15

2021 Synthesis Annexes: cas.cgiar.org/evaluation/publications/2021-Synthesis

Table of Acronyms

A4NH	CGIAR Research Program on Agriculture for Nutrition and Health
AAS	CGIAR Research Program on Aquatic Agricultural Systems
ACIAR	Australian Centre for International Agricultural Research
AfricaRice	Africa Rice Center
ARI	Advanced Research Institution
AWD	Alternate Wetting and Drying
BBSRC	Biotechnology and Biological Sciences Research Council
CARD	Coalition for African Rice Development
CapDev	Capacity Development
CAS	CGIAR Advisory Services Secretariat
CGRFA	(UN) Commission on Genetic Resources in Food and Agriculture
CIAT	International Center for Tropical Agriculture
CIRAD	Centre de coopération internationale en recherche
CSISA-MI	Cereal Systems Initiative for South Asia- Mechanization and Irrigation
CSIRO	Commonwealth Scientific and Industrial Research Organization
CSO	Civil Society Organization
DC	CGIAR Research Program on Dryland Cereals
EiB	Excellence in Breeding
FISH	CGIAR Research Program on Fish Agri-Food Systems
FP	Flagship Program
GL	CGIAR Research Program on Grain Legumes
GRIISP	CGIAR Research Program on Global Rice Science Partnership
ICAR	Indian Council of Agricultural Research
ICARDA	International Centre for Agricultural Research in Dry Areas
iDE	International Development Enterprises
IDO	Intermediate Development Outcome
IEA	CGIAR Independent Evaluation Arrangement
ILRI	International Livestock Research Institute
IRRI	International Rice Research Institute
ITPGRFA	International Treaty on Plant Genetic Resources in Food and Agriculture
JIRCAS	Japan International Research Center for Agricultural Sciences
L&F	CGIAR Research Program on Livestock and Fish
Livestock	CGIAR Research Program on Livestock Agr-food Systems
LSTHM	London School of Hygiene and Tropical Medicine
MAIZE	CGIAR Research Program on Maize

MARLO	Managing Agricultural Research for Learning and Outcomes
M&E	Monitoring and Evaluation
MEL	Monitoring, Evaluation, and Learning
MELIA	Monitoring, Evaluation, Learning, and Impact Assessment
NARES	National Agricultural Research and Extension System
NARS	National Agricultural Research System
NERICA	New Rice for Africa
NGO	Nongovernmental Organization
NRM	Natural Resource Management
OICR	Outcome Impact Case Report
PIM	CGIAR Research Program on Policies, Institutions, and Markets
PMU	Program Management Unit
PRMF	Performance and Results Management Framework (CGIAR)
QoR4D	Quality of Research for Development
RBM	Results-Based Management
R&D	Research and Development
R4D	Research for Development
RCT	Randomized Controlled Trials
RICE	CGIAR Research Program on Rice Agri-Food Systems
SIMEC	Strategic Information, Monitoring & Evaluation Committee
SLO	System-Level Outcome
SME	Subject Matter Expert
SO	CGIAR System Office
ToC	Theory of Change
ToR	Terms of Reference
USAID	United States Agency for International Development
WUR	Wageningen University & Research
VC	Value Chain
W1, W2, W3	Funding Windows 1, 2, and 3
WHEAT	CGIAR Research Program on Wheat
WLE	CGIAR Research Program on Water, Land, and Ecosystems

Executive Summary

1. Introduction and Methodology

The objective of this forward-looking synthesis was to bring together learning from a decade of experience with CGIAR research programs (CRPs), based on existing evaluative evidence. The purpose of this meta-review is to review lessons from the CRP experience to inform the development of future research programs of One CGIAR. The 2021 Synthesis and Lessons Learned from a Decade of CRPs is delivered in response to the request of the CGIAR System Council and aligned with the synthesis terms of reference endorsed by SIMEC in February 2021.

The synthesis examined evidence from the two phases of CRP implementation: 2011–2016 and 2017–2019. Four key issues were addressed: (1) patterns and trends between the two phases of CRPs related to the quality of science (QoS) and research for development, achievement of sustainable development outcomes, and management and governance; (2) systemwide issues affecting CRP achievements; (3) recommendations for the future orientation of CGIAR research and innovation; and (4) key evidence gaps and needs for future evaluations.

A narrative synthesis approach was used, employing secondary source data from 47 existing evaluations and reviews. External evaluations were systematically coded and analyzed by senior subject matter experts (SMEs) using a standardized analytical framework. A bibliometric trend analysis was carried out, and findings were triangulated against earlier syntheses and validated by members of the Independent Science for Development Council (ISDC), CRP leaders, and expert peer reviewers.

2. Trends and Lessons across the Two Phases of CRPs

The evaluations recognized the solid performance of the CRPs. Overall, CRPs were able to adopt radically new ways of working and integrate separate research efforts into coherent, focused programs, despite challenges that included a highly unstable funding environment and adaptations necessitated by the ongoing evolution of the CGIAR, its strategy, and results framework.

2.1 *In general, CRPs have:*

- Increased CRP and Center collaboration, creating new linkages and synergies between Centers and their researchers
- Increased the breadth, depth, and strength of external partnerships, particularly with eminent advanced research institutions (ARIs) and national innovation and policy systems
- Expanded relevant and innovative multidisciplinary research
- Opened new research areas for CGIAR, including nutrition and health, market chains, climate change, advanced genomics, and systems-based approaches, aided by new, innovative partnerships with ARIs
- Encouraged plant and animal breeding programs to move beyond productivity to address new challenges such as climate change, water scarcity, and nutritional quality
- Instituted new platforms in key areas, including GENE BANKS, breeding technologies, gender, and big data
- Confirmed the important role of CGIAR's agri-food research in informing policy discussions and change in traditional and newer areas such as climate change, natural resource management, resilience, and health and nutrition
- Mobilized partner capabilities within and beyond CGIAR to quickly respond to new challenges, such as fall army worm in maize

2.2 Quality of Science

Research inputs. Across both phases, evaluations found CRP research leaders and staff to be highly competent and productive; many are recognized leaders in their fields and represent Centers with long legacies. All CRPs were negatively affected by uncertain and declining Window 1/Window 2 (W1/W2) resources, especially in phase 2. By the 2020 reviews, W1/W2 funding accounted for only 22% of total CRP budgets, compared with 35% in 2014. Birner and Byerlee (2016) suggest that a minimum of 30–35% is required to effectively implement an integrative, collaborative cross-Center program.

Outputs. Outputs of high-impact, peer-reviewed scientific publications were high for most CRPs across both phases. Over time, publications in high impact-factor (IF) journals increased, owing in part to expanding partnerships with ARIs. H-indices were relatively low for several system CRPs (AAS, Drylands, Humidtropics), but evaluations noted the inadequacy of bibliometric analysis for research for development (R4D) research outputs.

Improved germplasm of varieties/strains, parental lines, and hybrids released to national agricultural research and extension systems (NARESs) were key outputs across CRPs. The research programs have significantly elevated CGIAR's focus on tolerance of major stresses, adaptation to climate change, and nutritional quality, in addition to productivity.

Challenges remain. Siloed approaches persisted between breeding and other disciplines and between Flagship Programs (FPs). However, PIM's FP1 collaboration with commodity CRPs strengthened the biophysical parameters of the IMPACT model and improved the flagship's capacity to use socioeconomic models. While cross-FP linkages remained generally weak, during phase 2 A4NH's Country Coordination and Engagement Unit improved FP coordination at the country level.

2.3 Quality of Research for Development

Credibility. One measure of QoS, bibliometric analysis of selected CRPs, showed a strong record of high-quality, peer-reviewed publications that improved across the two phases. However, science quality assurance processes, normally the responsibility of Centers, were inconsistently applied across CRPs.

Legitimacy. Ethics policies, a key measure of legitimacy, are also in Centers' domain. They were not enforced consistently across CRPs, although some (A4NH, PIM) remedied this shortcoming in phase 2.

Partnerships. Perceived legitimacy is critical to the effectiveness of external partnerships. CRPs prioritized active collaboration throughout the research process to ensure the relevance of outputs at the policy level (PIM, MAIZE, WHEAT, A4NH, FTA). WLE, AAS, Drylands, and Humidtropics all featured close engagement with communities, and Humidtropics developed multistakeholder processes to improve the relevance of social and technical innovations.

In general, evaluations found that CRPs did not focus adequately on the needs of the poorest and most vulnerable sectors and gave little attention to potentially valuable off-farm income and employment opportunities for rural women and youth. This gap also reflects limited social science capabilities and the CRPs' predominant focus on biophysical dimensions.

2.4 Delivery of Outcomes and Sustainability

CRPs delivered an array of outputs and outcomes at different levels. Most of the evaluations from which this synthesis was drawn focused primarily on scientific inputs, activities, and outputs; see the limitations section in Annex 3.

Commodity and systems CRPs. Commodity-focused CRPs applied high levels of scientific inputs to accelerate breeding cycles and achieve higher genetic gains. Sustainable intensification work has helped to optimize Genotype x Environment x Management for varieties/breeds under a wide range of conditions. The system-level CRPs using R4D approaches at the field level (e.g., participatory research, action research) faced significant challenges in delivering their intended outputs and outcomes because of unrealistic planning and expectations and/or because their areas of focus were insufficiently innovative for system transformation that would significantly benefit the poor.

Policy outputs and outcomes. By phase 2, evaluations indicated good progress on policy outputs and outcomes at different levels, particularly from the global integrating programs. These CRPs contributed significantly to policy discussions and change related to climate change, nutrition and health, response to economic shocks, and the sustainable management of land, water, and ecosystems.

Unclear assumptions about targets, timeframes, and CRP contributions. Across the CRPs, progress on outputs, outcomes, and scaling usually reflected programs' different maturity levels. Advanced programs had been established years before the CRPs (e.g., the Global Challenge Programs) and benefited from existing research work and partnerships. For newer programs, such as the system-level CRPs or A4NH FP1 on dietary diversity, careful assessment is needed to determine whether targets are realistic, measurable, and achievable. Hard consideration is also needed when targets include a last-mile delivery component, to ensure that clear and appropriate roles are set for CRPs and partners from the start.

The evaluations found current CGIAR metrics for development, sustainability, and resilience insufficient. Metrics on resilience, poverty alleviation, and sustainability are not user-friendly and little used, reflecting the lack of social science expertise. Progress indicators and impact assessment methodologies are particularly challenging for natural resource management (NRM) and systems research, where achieving impacts requires significant time, often exceeding the duration of a typical CRP.

Box 1. What is evaluative evidence related to system-level targets?

The Evaluation Synthesis questions sought to discover patterns and trends between the two phases of CRPs related to, inter alia, achievement of sustainable development outcomes, from the 30 CRP source reports the Synthesis team analyzed. The 2020 reviews projected that the CRPs would fall far short of the 2022 SLO and IDO targets (e.g. FTA, A4NH, MAIZE). Covid-19 was one factor projected to affect the achievement of SLOs and IDOs, together with reduction of phase 2 by one year across all CRPs. However, the much larger issue was the flawed process for establishing and measuring development outcomes. This hampered the ability of this Synthesis expert team, and past Review and Evaluation teams, to understand SLO and IDO achievements quantitatively and comprehensively. Introduced in Phase 2, outcome impact case reports (OICRs) highlight key achievements, without providing a quantifiable and comprehensive gauge of progress. This report provides recommendations on how to improve measurements and evaluability of future projects for accountability and learning.

"It is unrealistic that actual contributions FTA research makes to IDOs—as currently defined—can be monitored and aggregated quantitatively to yield FTA or CGIAR-level outcome measures. If forced, program staff is likely to creatively produce and report figures that will however be based on such uncertain critical assumptions that they are unlikely to reflect reality" (FTA 2014). "By year-end 2020 significant contributions had been made to all four end-of-program outcomes; it expects to contribute significantly to 2022 target outcomes but not fully reach them. Regarding FTA's 31 qualitative sub-IDO targets, FTA expects to fully meet expectations for 23 (74%) of these contributions" (FTA 2020).

"Although the outputs from this research potentially feed into RTB's adaptive and applied research on resilient cropping systems, most of this research (80–90 percent) is supported by other CRPs.... This gives the impression that limited research is being implemented on RTB crops that potentially will contribute to SLO3" (RTB 2015).

In phase 2, milestones were of limited use for showing progress toward IDOs and SLOs. *"Given that there are few measurable and specific targets for outcome milestones, it is difficult to precisely assess the extent of achievement of outcomes."* (MAIZE 2020).

Although FISH explicitly measured success in terms of achievements along the TOC pathways to SLOs, the numbers were noted to be below the projected outcomes.

The 2017 evaluation of results-based management (RBM) and 2020 CRP reviews concluded on difficulties in translating "SLOs [system-level outcomes] and IDOs into measurable results indicators and targets, partially due to diverse contexts behind indicators and targets of CRPs". Thus, although almost all CRPs accomplished an impressive array of outputs and outcomes, most CRP evaluations found it difficult to determine the linkages between the FP objectives, outcomes, and CGIAR-level SLOs. Evidence gaps were identified across a wide range of CRPs and themes that affected the assessment of their contributions to wider and more complex development outcomes.

Limited progress and coherence in the use of the theory of change (ToC) for management. In phase 2 most CRPs significantly improved their capacity to develop ToCs that align outputs with sub-IDOs, IDOs, and SLOs. However, beyond the initial planning of phase 2 research, CRPs rarely used TOCs to manage programs to deliver strategically toward defined IDOs and SLOs. The evaluations suggested a range of reasons, including the varying quality of TOCs, lack of familiarity with using TOCs in management, pressures to develop more user-friendly products to communicate progress and

achievements with partners and funders, and time constraints of research managers facing multiple, uncoordinated sets of reporting requirements.

WLE and systems CRPs emphasized the importance of convening processes in TOC development *"that can bring science and development together and engaging with diverse stakeholders to build mutual understanding and agreement on the potential of different options to manage landscapes and improve agriculture"* (WLE 2016). The skill set and mode of engagement needed to move from outputs to impact at scale is very different from the training and background of most CGIAR scientists.

Capacity development of uptake partners. Capacity at the country level is a key factor affecting scaling and sustainability of outcomes and impact. While individual capacity development for national scientists and partners featured strongly in many CRPs, all evaluations highlighted the need for a more strategic and systematic focus at the institutional level. Coherent capacity development strategies and processes to transfer research to national systems were largely lacking. More equitable partnerships are important so that all "own" and can commit themselves and their resources to the agenda.

Engaging the private sector. The CRPs have focused primarily on government, community, and ARI partnerships, but engaging the private sector will be critical for scaling innovations. Where CRPs did work with the private sector to extend innovations, there was no evidence of a private sector engagement strategy or analysis of the effectiveness and lessons learned from these efforts.

2.5 Management and Governance

Through the CRPs, cross-Center collaboration has now become a standard way of working. Major challenges remain, including funding, as discussed above, and the basis of its allocation between Centers. CRPs were usually highly dependent on, and benefited significantly from, the lead Center's well-established governance, finance, and human resources systems. However, in some CRPs (e.g., PIM, GL, DC) questions arose in phase 1 over the lead Center's dominance, specifically over transparency of funding allocation and potential conflicts of interest given the lead Center's role in governance as well as its fiduciary responsibility. New governance structures were agreed for phase 2 CRPs, including a new Independent Steering Committee (ISC). Phase 2 assessments noted the positive impact of the ISCs in improving trust and cooperation among researchers, synergies in research activities, and funding transparency, and in providing an independent source of advice.

There are also important management lessons from CCAFS, which developed a more independent structure, with the program management unit (PMU) outside the lead Center and expert non-CGIAR staff recruited for leadership positions. CCAFS's unique path bolstered its identity as a semi-independent integrative program and allowed it to focus on the strategic partnerships needed to achieve program goals, bringing fresh perspectives and increasing its capacity and the quality of its science (CCAFS 2020).

3. Conclusions: System-wide Issues That Have Strengthened or Weakened Achievement of CRP and CGIAR System Results

The continuing challenges faced by CRPs suggest the depth of change that is still required at the CGIAR System level, if One CGIAR is to contribute effectively to global food systems transformation.

3.1 CRP Funding and Its Implications

The 2009 CGIAR reform, which laid the ground for the CRPs, was premised on improving the efficiency and development impacts of CGIAR research. System funders aimed to move toward centralized and more stable funding, promote revenue growth, and bring common accountability for CGIAR performance through the CRPs and their intermediate development outcomes, against an agreed strategy and results framework. The reform also aimed to reduce the transaction costs of generating and managing several thousand individual projects across the Centers. A group of major funders initiated the CGIAR Fund, encouraging others to join in and grow the Fund as they gained confidence in central funding and reporting over time.

Through the CRPs, cross-Center collaboration became an established and accepted new way of working. In practice, however, the 2016 synthesis of CRP evaluations found that *"the consistent and wholesale shift away from W1/W2 funds toward W3 and bilateral funds threatens the fundamental objective of the reform"* and that *"there is a need to revisit the compact between the donors and the Centers in order to rebuild commitment to the CGIAR reform agenda and ensure the success of the second phase of CRPs."*

W1/W2 funding was intended to fuel innovative new collaborations and facilitate cross-Center portfolio management. Instead, CRPs were forced to rely increasingly on bilateral funding of Center-based projects driven by donor objectives that did not necessarily align with the CRP vision or provide longer-term funding to support more innovative research and new partnerships. Over phase 2, W1/W2 funding has declined to just over 20% of the total revenues, from 35% in 2014. A repeated finding across CRPs was that W1/W2 funding was insufficient to provide a central driver for the programs to deliver the overall purpose and intended outcomes. The shortfall in anticipated W1/2 funding and required use of W3/bilateral to sustain most of the programs has though carried significant implications for delivery of coherent CRPs.

Some Centers were able to use their reputations and fundraising prowess to successfully stitch together separate projects into more coordinated new areas of research, e.g., IFPRI's leadership of A4NH. But the research programs that were most severely affected by funding constraints during the two phases found it difficult to attract bilateral funding. In retrospect, they appear to represent a missed opportunity to advance understanding in areas that are highly relevant to CGIAR's ability to meet climate change and systems transformation challenges. These include long-term, place-based multidisciplinary research, such as the Sentinel Research Program of FTA and the systems approaches of Humidtropics, Dryland Systems, and AAS, none of which were granted a second phase.

Reasons for the shortfall and lack of donor confidence or ability to invest via W1/W2 have been identified in earlier CGIAR System-level discussions and were not evaluated in 2020. However, the evaluations included consultations with some specific donors. The evaluative evidence summarized in Box 2 explores the implications and impacts of the funding mechanisms used (i.e., W1/W2 and W3/bilateral) on delivery and success of the CRPs (see Box 2).

The core paradox recognized through these evaluations is that CGIAR occupies a unique role in development and is supported through development assistance funding through which donors look for development outcomes, yet by its nature CGIAR contributes to such outcomes through research and innovation products that also require engagement and commitment from many others to achieve sustainable development impacts. Two further challenges identified here are that the system lacks effective and convincing metrics for these interactions and that funding within agencies is not generally cross-linked to their bilateral development assistance to countries.

As called for in the 2016 synthesis report, a thorough re-examination of the basis and terms required for system donors to be able to commit to funding the new CGIAR initiatives, and the commitment of CGIAR to better demonstrate its contributions to development outcomes and impacts, remain fundamental. Many issues revolve around this interplay, requiring better understanding of national and community demand for innovations, recognition of the realities of innovation webs linking research to sustainable development outcomes, the need for a more holistic view of research *in* development processes themselves, and better coordination of donor support to these processes within and between agencies.

3.2 Reality Check: Time, Resources, Skills, and Partnerships Required for Scaling Development Outcomes

The CRPs have produced many high-quality and relevant research products, but there is a disconnect between the time and resources needed to achieve development outcomes at scale and the lifespan and typical research activities of a CRP. CRPs are required to articulate quantifiable IDOs and their contribution to the strategy and results framework. However, studies estimate that it takes 15 years or more for a successful pilot activity to reach a sustainable scale and be able to generate social and economic outcomes and impacts.¹

The IDOs cannot be achieved through CGIAR research alone. Sufficient prioritization and resources are required to develop relationships and strengthen the capacity of external public and private sector partners and stakeholders over time, starting from the beginning of the research process. Partners who perceive that they have an "ownership stake" in research that reflects their own priorities will be key collaborators in the sustainable scaling of research outcomes.

Over the past 50 years, the CGIAR has developed a deservedly strong reputation for its biophysical research. The CRP reviews pointed out that other disciplines and skills are essential, however, to create and maintain critical partnerships to inform research priorities and accelerate the progression from research development to development outcomes and impacts at scale. These include social science and

¹ Woltering et al. 2019, <https://www.sciencedirect.com/science/article/pii/S0308521X18314392>.

systems research in general, as well as investments to improve communication and joint work with public and private development partners.

Box 2. What is the evaluative evidence about CRP funding challenges?

Funding was a challenge in delivering all CRPs: After an initial boom following the initial reform and in response to the food price crisis, funding has steadily declined since 2014. Few donors have committed to multi-year central funding for programs as a whole, although the ones that have are significant. Funding has been unpredictable, with a frequent lack of clarity about donor commitments at the time of budgeting. Receipt of income was delayed due to both inconsistent receipt from donors in some cases due to delays in reporting, but often due to the misalignment of CRP annual report submissions and donor planning cycles, resulting in donors not having necessary information on hand to make timely allocation decisions. Budget volatility and the influence of donor funding affected program priorities, and cuts have taken a toll on partnerships, but the handling of budget cuts has generally been effective.

Donors expressed concerns about investing via W1/W2: Donors' concerns included lags between research activities and development outcomes, which make it difficult for donors to make required reports back to their own governments. Other donors claimed that research is "not sufficiently focused on results" and "too focused on publications" and mentioned the "need to show progress to stakeholders." Bilateral donors have been reluctant to support core infrastructure through project grants.

Expectations of CRPs and donors can be unclear: More needs to be done in developing partnerships with donors, with clear expectations on both sides and appropriate funding vehicles mobilized beyond research support. Development assistance funding often needs to show tangible results on the ground, but this can pressure CRPs to incorporate development outcomes into their ToCs that are unrealistic, given the many steps and actors involved between researching a new technology and its adoption by farmers. Success can be enhanced through partnerships and engaging with potential funders from the outset, matching CRP aims with donors' specific priorities to ensure that all required innovation system needs are addressed. Early engagement and dialogue with governments and donors are key to supporting wider uptake and impacts.

Rising share of bilateral funding has implications: CRP managers have direct responsibility for only W1/W2 funds—around 20 percent of program budgets. They have had to secure bilateral projects, coherently fitting these to their CRP frame. Formulating and negotiating many small projects brought high transaction costs. As reliance on W3/bilateral funding increased, priority setting shifted. Strong influence from a donor can be positive or negative, but many bilateral projects are designed outside a CRP, with donors pursuing their own aims and determining what is done and how it is implemented. In some cases, successfully leveraging donor resources swayed program focus away from the program's own vision and ToC and emphasized other actions and countries. Cross-cutting issues such as monitoring, evaluation, learning, and impact assessment (MELIA) and gender became vulnerable when W1/W2 funds were scarce.

Balance between research and development should be clarified: Donor emphasis on short-term results and field impacts can take work downstream into technology delivery and extension even where CGIAR has no comparative advantage. Scaling up innovations can address grand challenges but requires a significant shift in orientation and strategies away from basic research. CRPs are faced with hard choices: should they invest in upstream research or focus more on delivery? The role and positioning of CGIAR in regard to other actors needs to be better clarified and understood all around.

Multiple reporting lines impose inefficiencies: Even with a more effective monitoring and evaluation system, tailored reports would still need to be generated to suit diverse demands of W1/W2 and W3/bilateral donors. Researchers leading W3/bilateral-funded projects prioritized reporting to their own donors, using donors' reporting formats and timelines. Managing agricultural research for learning and outcomes (MARLO) platform is essential for CRP reporting, but for many it comes second to project-level reporting to donors. A donor-imposed system of milestone reporting was noted to have a highly negative effect, wasting time and undermining the quality of the preferred ToC-based reporting.

Many NGOs and technical organizations funded by bilateral and multilateral agencies that also fund CGIAR already have established relationships with public and private partners at the national and local level. With facilitation from funding agencies, CGIAR could productively tap these relationships to avoid

duplication and reduce the need for CGIAR to develop the skills and relationships to foster “last-mile” partnerships on its own. One example of such an “unusual partnership” is the Cereal Systems Initiative for South Asia- Mechanization and Irrigation (CSISA-MI) program in Bangladesh, which featured an alliance between CIMMYT, International Development Enterprises (iDE), and local private sector entrepreneurs that resulted in a widely adopted model for agricultural machinery service provision. Challenges to full scaling remain, including limited access to finance for agricultural machinery service providers. However, the CSISA-MI experience, and other innovative programs such as the CIMMYT–Government of Mexico MasAgro collaboration, signal the considerable potential for better linking CGIAR with development partners.

3.3 Inadequate Monitoring and Learning for Development Outcomes

A major concern identified by the CRP assessments is the inadequacy of CGIAR monitoring, evaluation, learning, and impact assessment (MELIA) systems for planning, monitoring, and incorporating learning from CGIAR R4D activities.

Across the CRPs, inputs, outputs, and their scientific quality have been well measured, but the same is not true of CGIAR’s contribution to wider and more complex development outcomes. For the CRPs more generally, the continuing focus on traditional measures of research success rather than harder-to-quantify measures—e.g., resilience, quality of partnership development, and attention to capacity development—also reflect the lack of social science expertise in CGIAR today. This incentivizes researchers toward the development of technologies and innovations, with less consideration of the pathways through which stakeholders engage, inform, and take ownership of research agendas. With inadequate resources, capacities and attention to the engagement of others in priority setting and uptake pathways, researchers’ efforts to propel research outputs and outcomes to widespread adoption is limited.

The ToCs and definition of impact pathways were a step forward in phase 2 CRPs, but the fact that so few CRPs were able to use the TOCs and related measures for practical, adaptive management purposes shows there is still a long way to go.

4. Actionable Recommendations

These recommendations are derived from evidence-based conclusions on issues found to be common across many or all CRPs, and many draw on still-relevant recommendations from previous evaluations.

4.1 Recommendations for One CGIAR

- 1** Invest in preserving and taking forward valued elements developed through the CRPs: infrastructure, relationships, processes, tools, and innovations.
- 2** Prioritize partnership development and stakeholder engagement. Develop and implement a systemwide strategy for equitable engagement and effective communication with partners and stakeholders of all categories in the foresight, planning, delivery, and follow-through of CGIAR research, with metrics derived from partner perspectives.
- 3** Focus much more on institutional capacity development, especially of national “boundary” partners. Develop and implement a systemwide strategy and partnerships with other agencies to facilitate development of required capacities for uptake, transformation, and use of CGIAR products. To help achieve development outcomes, CGIAR and its programs should more actively advocate and help leverage financial resources for capacity development of national partners in pathways to impact. This requires a clearer and more consistent positioning of CGIAR and its role within the R4D continuum and new metrics on the efficacy of capacity development in enabling others to take forward CGIAR’s research processes and products for themselves.
- 4** Define CGIAR’s comparative advantage in delivery of different elements of the ambitious 2030 Research and Innovation Strategy and its projected scale of funding: review where internal investments and capacities are most needed and where gaps can be more effectively met through external partnerships.
- 5** Strengthen country and regional coordination structures to enable all CGIAR Centers and research initiatives to explore integrative solutions at local, landscape, and relevant subnational, national, and regional scales, ensuring coherent and responsive engagement with national

stakeholders and agendas. These can leverage the assets and scientific knowledge, local relationships, and reputation developed by Centers over five decades.

- 6 Operationalize a high-quality, common approach to research ethics and science quality and their measurement.
- 7 Maintain effective knowledge management to track processes and findings through successive phases of work and maintain public access to key CGIAR documents and research data.
- 8 Shift practices and evaluation away from seeking to attribute development impacts to CGIAR research and toward determining and valuing the essential contribution CGIAR is making with others, both through its research and by mobilizing collective actions among diverse public, private, and civil society partners to transform innovation systems for development impact.
- 9 Enhance determination of QoS through bibliometric analyses, and facilitate comparison across CRPs and new research initiatives by (1) maintaining the same data sources over time, (2) obtaining citation data annually to enable direct comparisons unaffected by the number of years elapsing, (3) retaining data from analyses in their raw format, including all metadata, to allow data to be reanalysed in the future and visualized in new ways, and (4) developing standard guidance and indicator definitions.

4.2 Recommendations for the Three Strategic Action Areas (AA)

4.2.1 Systems Transformation (AA1)

- 10 Rather than tackling climate change, NRM, and agriculture for nutrition and health separately, CGIAR should consider them together, holistically, exploring science-policy synergies and trade-offs across the areas as food systems transform.
- 11 To achieve a stronger focus on poverty reduction across all programs, target the rural resource-poor, women, and those most disadvantaged. Increase attention to understanding and addressing the equity impacts of policies, shocks, and risks faced by poor people in taking up technologies and research solutions.
- 12 Address the linkages between environmental sustainability and resilient agri-food systems. Relationships between the dynamics of environment, ecosystems, biodiversity, and livelihoods in agro-ecosystems will require significant attention.
- 13 Identify a handful of place-based programs in high-priority agro-ecologies, where the triple challenge of achieving sustainable food production, enhancing human well-being, and conserving ecosystem services can be addressed and where national commitments bring opportunity for impact at scale through integrated innovation systems.

4.2.1 Resilient Agri-Food Systems (AA2)

- 14 Reorient work to focus more on the vulnerable poor, in particular women and the disadvantaged and those at greatest risk from natural resource depletion, severe climate change impacts, economic deprivation, and conflicts.
- 15 Improve assessment and metrics related to risk and resilience, and co-develop social and technical innovations with at-risk populations.
- 16 Foster adoption of technical and social innovations at scale, as required to achieve system transformation, and give greater emphasis to research on scaling science and implementation science.

4.2.2 Genetic Innovation (AA3)

- 17 Ensure that high priority is given to nutrition, health, resilience, and environmental sustainability objectives in research groups focused on genetics.
- 18 Increase inclusiveness in defining product profiles, executing programs, and delivering outputs, to better contextualize variety development and tailor research to diverse agricultural communities and to the needs of children, youth, women, and other at-risk or marginalized groups.
- 19 Prioritize seed sector development, including by expanding partnerships with the private sector and civil society and strengthening key policies and regulations.

- 20** Catalyze partnerships with other research and innovation partners in defined systems to enable crop system diversification and improve access to affordable, healthy diets.
- 21** Accelerate the modernization and technical capacity development of plant-breeding programs across Centers and in national program partners.
- 22** Integrate research with wider development and investment commitments related to climate change adaptation and mitigation.
- 23** Engage strategically with policies (e.g., ITPGRFA, CGRFA) around the value of germplasm diversity, farmers' and breeders' rights to plant and animal genetic resources, and international transfer agreements to ensure access to and availability of diverse and valuable germplasm, improved varieties and strains, and crop wild relatives.

4.3 Seven Ways of Working

4.3.1 Embracing a Systems Transformation Approach, Seeking Multiple Benefits Across Impact Areas

- 24** Ensure that public, private, and civil society stakeholders are involved in foresight and priority-setting processes and have a sense of ownership about the research agenda.
- 25** Strengthen the systematic incorporation of equity issues into research design and analysis. Diversify partners and skills—including, for example, social scientists and experts from the private sector, sustainable finance, and humanitarian sectors—to better address the root causes of sustainable development challenges. Expand socioeconomic work, including poverty and livelihood assessments, adoption studies, policy and institutional analyses, and in-depth gender and youth studies, with strengthened in-house capacity and/or additional partners.
- 26** Invest in training researchers in systems science. Build research from a shared understanding of food systems that integrates objectives related to production, livelihoods, environment and biodiversity, and health and nutrition; that takes a holistic approach to agri-food systems and risk management; and that uses participatory innovation approaches to engage with farmers and rural communities.
- 27** Strengthen MELIA metrics, and develop user-friendly, streamlined reporting systems based on simple, nested ToCs—developed with and owned by partners and stakeholders—that enable required baselines, actions, capacities, and responsibilities to be coherently planned in pursuit of desired outcomes.
- 28** Tailor corresponding metrics to CGIAR's comparative advantage and realistic expectations of CGIAR's contribution to sustainable development outcomes across the 5 impact areas.
- 29** Incentivize the use of MELIA metrics for progressive cycles of evidence-based learning and adaptive management, working in close collaboration with partners and stakeholders, to optimize delivery and impacts. Increase the use of mixed-method designs in evaluations, with metrics for outcome pathways that go beyond CGIAR and its immediate boundary partners.
- 30** Improve the coverage of cross-cutting themes (e.g., gender, youth) in MELIA by strengthening evaluators' relevant disciplinary skills as applied to evaluation design and implementation.
- 31** Expand the availability of technical assistance on MELIA to research managers, scientists, and partners.

4.3.2 Leveraging Ambitious Partnerships for Change

- 32** Develop strategies for developing partnerships and institutional capacity, to facilitate a more systematic approach in both areas. Establish explicit time-bound targets and exit strategies for the progressive transfer of responsibilities and resources to enable local partners to sustainably take on a research or innovation area for themselves.
- 33** Draw on CGIAR's value as a broker of networked actions by making greater use of research and development partnerships to fill knowledge and skill gaps in research processes and innovation webs, enabling CGIAR to focus on its own strengths and areas of comparative advantage. These partnerships, including South-South partnerships, should include the private sector throughout the food system, non-CGIAR ARIs, small and medium-sized enterprises, and civil society organizations (CSOs), to help scaleup innovations, value addition, and market access. Facilitate

partnerships linking non-CGIAR ARIs to local and national partners for collaborative research and capacity development in new initiatives. Explore opportunities for CGIAR programs to contribute productively to national development agendas, foster synergies, and reduce duplication of effort. For example, the GENE BANK and Excellence in Breeding (EiB) platforms were established as service providers to CGIAR but have the potential to strengthen genetic conservation and use and advanced breeding capabilities in national systems.

4.3.3 Positioning Regions, Countries, and Landscapes as Central Dimensions

- 34** Put higher priority on ensuring that research agendas respond to local, national, and regional strategies and initiatives to facilitate the achievement of outcomes at scale. Initiate or strengthen long-term, transdisciplinary research at dedicated field facilities strategically located in relevant landscapes of developing countries. Co-locate activities from many programs in these geographic areas to better coordinate outcome-driven research activities, build partnerships, and share infrastructure.
- 35** Develop consistent policies and practical, ethical guidance to inform CGIAR engagement with local partners at different levels (communities, government, private sector, NGOs, ARIs). Communicating in the right way with local partners is essential; CGIAR should expand its in-house communications and outreach capacities and ensure that country-based staff are well trained. Develop guidelines for future work based on the experiences of the systems CRPs and Global Integrating Programs in developing, funding, and managing Platform-based research initiatives with broadening participation and community engagement.

4.3.4 Generating Scientific Evidence on Multiple Transformation Pathways

- 36** Strengthen social science capacities by increasing in-house resources and/or making better use of skilled external partners. Integrate social scientists into action research projects, and develop appropriate incentives to encourage interdisciplinary and systems research.
- 37** Invest in creating a shared vision—including stakeholders and researchers—on what could be achieved in a group of research activities at the region, country, landscape, or community level and a ToC on how to achieve change. A successful process will require significant attention to facilitating communications among the different levels of researchers and stakeholders.

4.3.5 Targeting Risk Management and Resilience as Critical Qualities

- 38** Expand work on assessing risk and resilience and managing risk throughout the food system by strengthening CGIAR capacities or engaging external partners. Put a higher priority on improving resilience to climate and pest stresses when developing, adapting, and assessing technologies and innovations for crops and livestock.

4.3.6 Harnessing Innovative Finance

- 39** Collaborate with ARIs and the private sector on action research that unlocks access to finance, inputs, and innovation-based enterprise opportunities for women, youth, and other marginalized groups, building on index insurance, blended (public-private and public-private-producer) finance models, and other emerging approaches.
- 40** Pursue direct links between CGIAR R4D actions—coordinated in country—and official development assistance (ODA) loans and grants to countries, as well as direct co-financing through such mechanisms where feasible and where demanded by national programs.

4.3.7 Making the Digital Revolution Central to Our Way of Working

- 41** A wholesale review of CGIAR capacities and opportunities around big data and practical field applications for pro-poor sustainable development should involve:
 - Expanding the use of remote sensing and GIS;
 - Exploring ethical applications of artificial intelligence, big data, and citizen science that would specifically benefit the poor;
 - Assessing lessons from the rapidly expanding use of open data and digital tools for breeding, weather and agronomic information, extension, and marketing.

Action on all recommendations would increase the contribution of CGIAR research towards sustainable development impacts. Selected recommendations are particularly well aligned towards the five CGIAR Impact Areas, as seen in the box below.

Box 3. How do Recommendations Map to Impact Areas

Impact Area 1: Nutrition, health and food security

An integrated approach is needed, recognizing local contexts and the synergies and trade-offs between different dimensions of sustainable agri-food development. CGIAR should catalyze partnerships with other research and innovation partners in defined systems to enable crop system diversification and improved access to affordable, healthy diets and ensure that high priority is given to nutrition, health, resilience, and environmental sustainability objectives in research groups focused on genetics.

Impact Area 2: Poverty reduction, livelihoods, and jobs

Program planning and implementation should adopt holistic and integrated approaches to agri-food innovation as an engine of poverty reduction and sustainable rural development, linked to national agenda and engaging diverse partners to turn innovations into economic opportunity for the resource poor. To that end, prioritize partnership development and stakeholder engagement and develop and implement a systemwide strategy for equitable engagement and effective communication with partners and stakeholders of all categories in the foresight, planning, delivery, and follow-through of CGIAR research, including:

- Strengthen social science capacities through recruitment and partnership and focuses much more on institutional capacity development, especially of national “boundary” partners.
- Develop and implement a systemwide strategy and partnerships to facilitate development of required capacities for uptake, transformation, and use of CGIAR products.
- Increase advocacy for and help leverage financial resources for capacity development of national partners in pathways to impact.
- Foster adoption of technical and social innovations at scale, to achieve system transformation

This requires a clearer and more consistent positioning of CGIAR and its role within the R4D continuum and new metrics on the efficacy of capacity development in enabling others to take forward CGIAR’s research processes and products for themselves.

Impact Area 3: Gender equality, youth, and social inclusion

An increased focus is required on inclusive planning and delivery for innovation to provide both on and off-farm economic opportunity from agri-food innovation, reduce conflicts and pressures for out-migration and more directly address the needs of resource-poor smallholders, rural women and youth, indigenous peoples and other disadvantaged groups

Impact Area 4: Climate adaptation and mitigation

Principles generated through CRPs should be brought into future research to transform food systems and bring emphasis, alongside climate smart agriculture, on greenhouse gas mitigation and adaptation measures for mixed smallholdings. Integrate research with wider development and investment commitments related to climate change adaptation and mitigation. Rather than tackling climate change, NRM, and agriculture for nutrition/health separately in CGIAR, target research to science-policy engagement as part of a broader, integrating effort on transforming food systems.

Impact Area 5: Environmental health and biodiversity

To build on good immediate outcomes on NRM management, recommend a more integrated approach to links between environmental sustainability and resilient agri-food systems, ideally with place-based programs in high priority agro-ecologies, where the triple challenge of achieving sustainable food production, enhancing human well-being, and conserving ecosystem services can be addressed.

5. Gaps and Needs for Future Evaluations

The systematic review of evaluative evidence identified themes, where evidence was too limited, fragmented, or inconsistent to allow us to make conclusive findings about their worth.

5.1 Thematic Gaps

- A. **Integration of cross-cutting issues in CRPs:** Beyond gender dimensions, little evidence was found from which to draw conclusions about other groups of concern such as youth.
- B. **Capacity development outcomes:** While the evaluations largely covered immediate capacity development activities and outputs, little evaluative evidence was available by which to determine their value and resulting outcomes.
- C. **Challenges and opportunities related to digital innovations:** Evidence on digital innovation in the CRPs was scant, as this was not a subject of evaluation.
- D. **Stakeholders' engagement and CGIAR positioning within the AR4D continuum:** Evidence was fragmented and not consistent enough to assess the quality of CRPs' engagement with national public and private partners.
- E. **CRP outcomes, impacts, and sustainability:** Evaluation reports recognized the need for better assessment and documentation of outcomes and impact. The lack of appropriate monitoring, evaluation, and learning (MEL) with regard to CGIAR's development objectives is a significant handicap in seeking to demonstrate the extent to which CRPs contribute to outcomes and in making recommendations for the future.
- F. **Lack of evaluation of outcomes:** The lack of follow-through in evaluating outcome measures—such as the use made of capacity development; changes in practices, policies, and institutions; and direct and indirect benefits in CGIAR's "sphere of influence"—is particularly surprising, given the premise upon which the CRPs were established and funded.

5.1 Methodological Gaps

- G. **Context of evaluations:** The CGIAR Independent Evaluation Arrangement (IEA) and the CGIAR Advisory Services Secretariat (CAS) did not generally undertake country and regional evaluations themselves. Moreover, local stakeholders' voices were largely absent from 2020 reviews.
- H. **Metrics on CGIAR mission:** The dearth of metrics tailored to CGIAR's mission reduced the value and consistency of evidence gathered and analyzed in the evaluations.
- I. **Cross-cutting themes:** There was a lack of depth in assessing cross-cutting themes—e.g., gender, youth, climate change, NRM, and capacity development outcomes—in CRP evaluations.
- J. **Phase 2 assessments:** The lack of second phase assessments of the Gender and GENE BANK Platforms constrained the comprehensive analysis of progress along and between phases.
- K. **Contribution analysis:** Methods remain skewed towards attributing results and outcomes to CGIAR's work. Given that CGIAR is only one actor within the complex reality of innovation webs, there is a strong need for more contribution analysis, using mixed-method approaches to determine, from the data and perspectives of others, the scale and value of the contribution made by CGIAR's work toward achieving sustainable development outcomes.

1 Background and Synthesis Context

In 2018 CGIAR embarked on an ambitious reform, One CGIAR, across the entire CGIAR system. To be launched in 2022, the reform aims to bring a unified and compelling mission, unified governance, institutional integration, and a new research modality as well as to attract more, and pooled, funding. The CGIAR 2030 Research and Innovation Strategy² proposes to deploy scientific innovations at scale for greater impact where it is most needed. CGIAR, a global partnership implementing agricultural research for a food-secure future, currently delivers its portfolio through CGIAR Research Programs (CRPs).³

The CGIAR Advisory Services (CAS) Secretariat supports and facilitates CGIAR's independent advisory services, comprising the Independent Science for Development Council (ISDC), the Standing Panel on Impact Assessment (SPIA), and an independent evaluation function. In 2020 the CAS Secretariat, through its evaluation function, completed independent reviews of the twelve (12) CRPs. Prior to this, the CGIAR Independent Evaluation Arrangement (IEA) (predecessor to the Evaluation function in CAS) coordinated CRP evaluations between 2014 and 2016 and subsequently produced a synthesis of CGIAR reviews and evaluations⁴ covering the timeframe 2010 to 2016, the first phase of CRPs. The individual CRP evaluations provided evidence for the second phase of CRPs, 2017–2021.

With the completion of the 2020 CRP reviews and the launch of a new research modality to advance the One CGIAR research and innovation strategy, a synthesis of evidence on evaluative lessons learned is timely. At the request of the CGIAR System Council, the evaluation function under the CAS Secretariat has sought to synthesize evidence from almost exclusively independent, external evaluations and evaluative reviews conducted over the two phases, between 2014 and 2020. This evaluation synthesis aggregates and crystallizes findings, lessons learned, and recommendations from the independent evaluations and reviews conducted over this period. It provides information needed to evidence-fit and help shape the future research portfolio and emerging investment plan and aims to serve the dual purpose of accountability to CGIAR funders and learning. Primary intended users are the CGIAR System Council and System Board, the CGIAR Executive Management Team, and the System Office as a reference for system management in the change process. A second group of users consists of investors beyond the System Council as well as teams designing new initiatives under One CGIAR.

1.1 CGIAR Reform and Evolution of the CRPs

The CGIAR reform of 2008–09 addressed systemic challenges faced by the CGIAR system. CGIAR was viewed as falling short of its potential and risking a rapid loss of relevance. Funding had plateaued during a critical world food price crisis, threatening CGIAR's effectiveness and the contribution of international public goods to agricultural innovation. Key reforms were set to deliver the CGIAR's vision: *"To reduce poverty and hunger, improve human health and nutrition, and enhance ecosystem resilience through high-quality international agricultural research, partnership and leadership."* CGIAR investments aimed to become results focused and timebound through a Strategy and Results Framework (SRF), with all channels of funding and defining how CGIAR aimed to achieve its strategic objectives.

In the 2008–09 reform, the Centers and all CGIAR funders, with their respective partners, took on shared responsibility and mutual accountability for managing toward outcomes. A key intention was to move away from a fragmented research agenda toward a programmatic approach, with much larger programs funded for longer periods, based on their expected outputs and outcomes. This far-reaching reform resulted, in 2011, in the commissioning of the first of a new set of 15 large and ambitious multi-Center CGIAR Research Programs (CRPs), driven by their potential impact on development. Partnerships were considered essential for designing research programs and delivering research outcomes and impact, so incentives for partnership through the Centers were to be built into all levels of the CGIAR System. The CRPs aimed to reduce duplication and present evidence for the relevance, scientific value, efficiency, and effectiveness of the CRPs and Centers in delivering greater development

² CGIAR, CGIAR 2030 Research and Innovation Strategy, 2020, <https://cgspace.cgiar.org/bitstream/handle/10568/110918/OneCGIAR-Strategy.pdf>.

³ CGIAR, Research Portfolio 2017–2021, <https://www.cgiar.org/research/research-portfolio/>.

⁴ R. Birner and D. Byerlee, *Synthesis and Lessons Learned from 15 CRP Evaluations* (Rome: Independent Evaluation Arrangement of CGIAR, 2016), <https://cas.cgiar.org/sites/default/files/pdf/Synthesis2016web.pdf>.

outcomes. Adopting an agricultural research-for-development (R4D) approach throughout meant that all research priorities and activities became guided by their potential contributions to system-level outcomes (SLOs).⁵

The emergence of a more coherent programmatic approach required a deep cultural change from a Center-focused research basis to that of CGIAR Research Programs. This shift in turn required new research governance and matrix management in program implementation, to harmonize and bring coherence to the whole portfolio. In particular it required coordinating the management and use of central funds invested through the CGIAR Fund (Windows 1 or 2 [W1 or W2], depending on the degree of direction) and those obtained from Fund donors for specific purposes (W3) or bilaterally through direct commissioning of a Center's work (bilateral).

The CRPs have been implemented in two phases: 2011–16 and 2017–21. The initial 2011 CRP proposals were developed in parallel and individually, rather than as a portfolio.⁶ The first of these built on previously existing initiatives, whether legacy research or the Global Challenge Program initiatives on Water & Food (2002) and on Climate Change, Agriculture & Food Security (2008). Less attention was given to CRP portfolio issues such as managing inter-CRP linkages and developing harmonized operational plans, governance structures, and a performance management framework.⁷ During the first phase, the 15 CRPs underwent external review (Annex 4). As shown in Figure 1, significant changes were made to the portfolio: several Systems-research-based programs were terminated, while other new programs and cross-cutting platforms were initiated to address important common needs.

The Action plan of the System's 2012 SRF was intended to reflect demands from national and regional partners for results on the ground with potential for out-scaling, and the need to produce international public goods (IPGs) within the agricultural R4D institutional landscape. The System defined specific CRP targets as measurable intermediate development outcomes (IDOs) toward bigger development goals. The IDOs aimed to provide measures against which contributions by CRPs to any given outcome could be assessed and the significance of CRP outputs for progress toward these outcomes be understood.

"It is necessary, but not sufficient, for research to generate high quality research outputs. While it is clear that impacts can rarely be attributed to research alone, as many other actions and actors contribute to achieving impact, the art of designing meaningful IDOs is to find the middle ground between outcomes that are a reasonable proxy for impact and outcomes that researchers can be held accountable or responsible for. While researchers cannot achieve IDOs alone, they can be held accountable for developing partnerships that do achieve outcomes, and they can be jointly responsible for achieving outcomes with their partners."

The CGIAR Strategy and Results Framework Action Plan (2016-2030), October 2012

CGIAR governance and management also evolved considerably over the period, with a further structural reform following the 2015 Mid-term Review of CGIAR,⁸ which resulted in additional changes to the System's governance and management structure. The review recommended revising the SRF to enable prioritization and adequate resourcing of research: "The Research targets and indicators in the SRF should guide individual CRPs and facilitate effective periodic review of research progress of the CRP portfolio, to determine whether the portfolio is achieving the broad strategic objectives." The CGIAR Strategy and Results Framework 2016–2030⁹ hence set out SLOs and associated IDOs, with aspirational development targets for each, recognizing that for success, actions must align with specific national commitments or agreed regional objectives.

⁵ CGIAR Consortium Office, [The CGIAR Strategy and Results Framework Action Plan](#), October 2012.

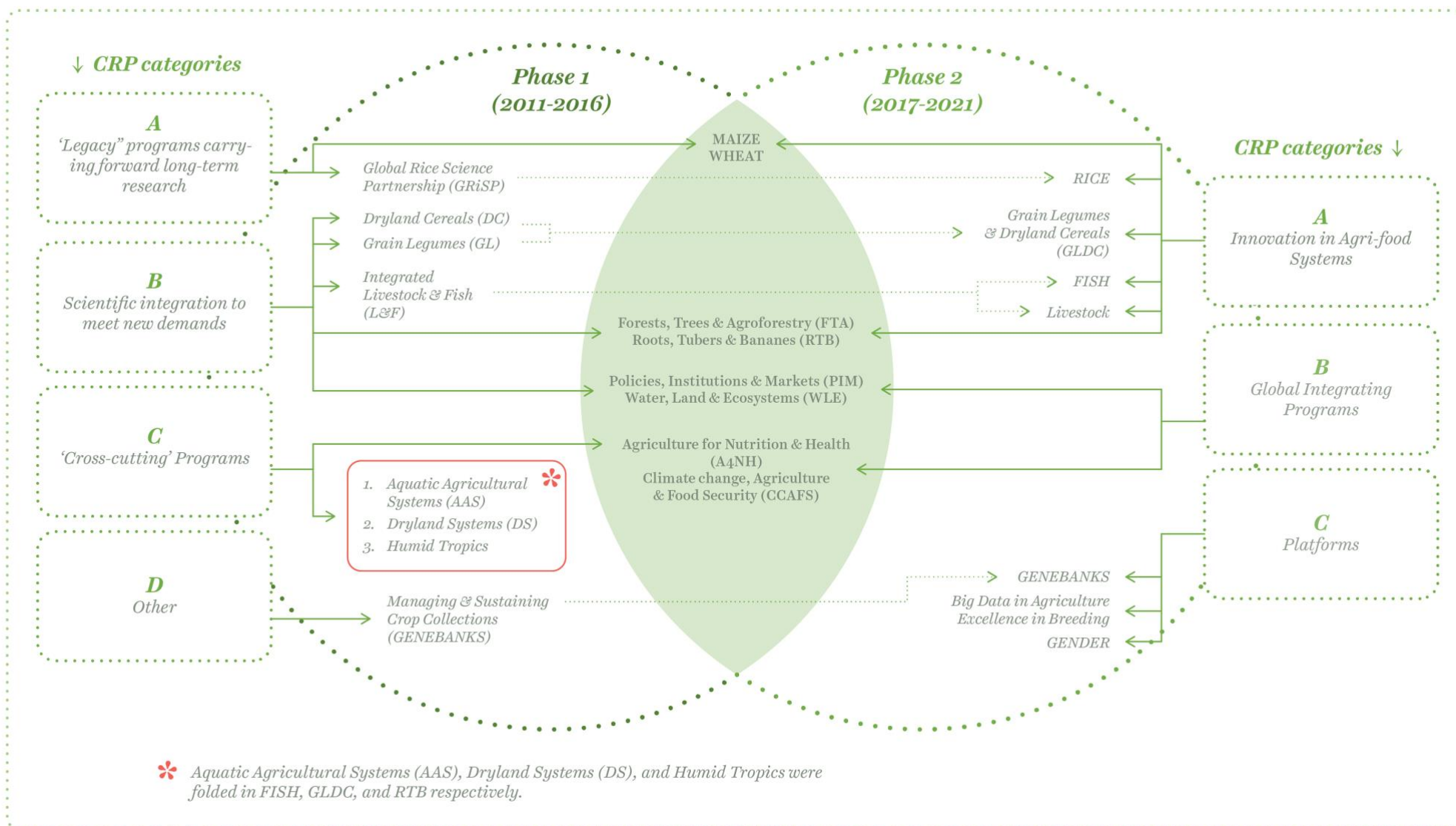
⁶ CGIAR, [The Integrated CGIAR Reform Proposal](#), 2008.

⁷ CGIAR Centers, [Delivering on CGIAR Strategy and Results Framework, CRP II Portfolio](#), 19 May 2015.

⁸ Rachel Kyte, *Final Report from the Mid-Term Review Panel of the CGIAR Reform (English)* (Washington, DC: World Bank, 2016), <http://documents.worldbank.org/curated/en/924981467998765166/Final-report-from-the-mid-term-review-panel-of-the-CGIAR-reform>.

⁹ CGIAR Consortium Office, *CGIAR Strategy and Results Framework, 2016–2030*, 2015, <https://cgspace.cgiar.org/handle/10947/3865>.

Figure 1. Evolution of the CRPs over two phases of implementation

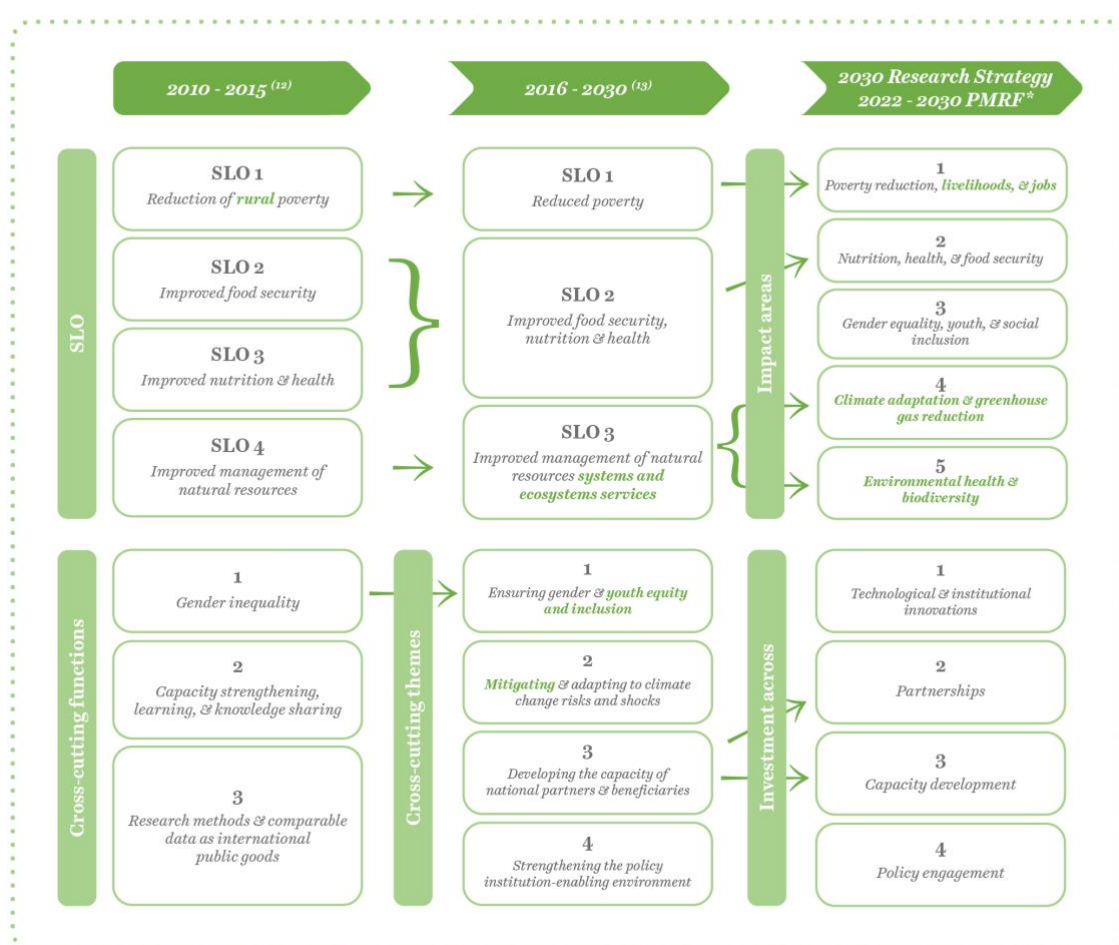


The second-phase CRP portfolio built on the 2010–16 CRP portfolio but differed in that it aimed¹⁰ to, among other things:

- explicitly support innovation in agri-food systems (AFSs) and integrated approaches to achieve development outcomes;
- clearly define theories of change (ToCs), set forth impact pathways, and identify outputs, research outcomes, and specified targets toward IDOs and sub-IDOs, from as defined in CGIAR’s 2016–30 SRF;
- make explicit collaboration among the AFS CRPs in defined geographies, with well-defined coordination and partnership mechanisms, led by a specified Center or CRP in each country or site¹¹;
- have a holistic approach to capacity development.¹²

Details of Center participation in the two phases of CRPs and the CRP objectives of phase 2, with information on flagship programs in both phases, are summarized in Annex 2.

Figure 2. Mapping of system-level outcomes and cross-cutting themes by SRFs



The 2016–30 SRF remains valid in One CGIAR under the new Research Strategy and its 2022-2030 Performance and Results Management Framework (PRMF) with envisaged adjustments to cross-reference the redefined impact areas.

¹⁰ CGIAR, [CGIAR Research Program Portfolio \(CRP2\)](#) (2015).

¹¹ CGIAR Consortium Office, [Update and Guidance on Site Integration](#), 2016.

¹² CGIAR Consortium Office, CGIAR Research Program Portfolio Report 2014, https://cgspace.cgiar.org/bitstream/handle/10947/4086/CGIAR-CRP-PortfolioReport_2014.pdf?sequence=1

1.2 Objectives and Questions

The primary purpose of the 2021 synthesis is to produce high-level findings, learning, and recommendations, generated from evaluative evidence within CGIAR (from 2014 to December 2020), with the following four specific objectives:

A. Identify trends and patterns around selected priority topics, such as quality of science, progress toward outputs, performance (achievement of objectives), CRP implementation and management, and future orientation (in One CGIAR).

B. Provide evidence-based recommendations on future orientation aligned with One CGIAR's:

- *Three interlinked strategic action areas (AAs)*: systems transformation, resilient agri-food systems, and genetic innovation
- *Five impact areas (IAs)*: (1) nutrition, health, and food security, (2) poverty reduction, livelihoods, and jobs, (3) gender equality, youth, and social inclusion, (4) climate adaptation and mitigation; and (5) environmental health and biodiversity
- *Seven ways of working*: (1) embracing a systems transformation approach, seeking multiple benefits across impact areas; (2) leveraging ambitious partnerships for change; (3) positioning regions, countries, and landscapes as central dimensions; (4) generating scientific evidence on multiple transformation pathways; (5) targeting risk management and resilience as critical qualities; (6) harnessing innovative finance; and (7) making the digital revolution central to our way of working.

C. Propose mechanisms for uptake and application of recommendations and lessons learned moving forward toward One CGIAR.

D. Ascertain the key evidence gaps and needs for future evaluations to provide direction to the creation of a multi-year evaluation plan (2022–24).

Accordingly, to inform future research, the 2021 synthesis was expected to bring to the fore operational and strategic evidence to answer four key questions:

1. What lessons can be learned between the two phases of CRPs in comparable areas of exploration in the CRP evaluations and reviews (i.e., effectiveness; quality of science (QoS); cross-cutting themes of gender, capacity development, youth, and climate change; as they apply to the newly defined R4D modalities of One CGIAR)?
2. What patterns and lessons from CGIAR System-wide issues have strengthened or weakened the achievement of CRP results and those of the CGIAR System?
3. What recommendations can be made on future orientation along the key priority themes, corresponding to One CGIAR's five impact areas, three interlinked strategic action areas, and seven ways of working?
4. What are the key evidence gaps and needs for future evaluations that must be addressed as CGIAR develops its multi-year evaluation plan (2022–24) and transition to One CGIAR?

1.3 Overall Approach

The overall approach is summative and formative. A predominately qualitative mixed-methods design was implemented. "*Narrative synthesis*" was selected and used as an approach to the systematic review and synthesis of findings from evaluation reports and reviews. This approach relies primarily on the use of words and text to summarize and explain the findings¹³. The data were aggregated and structured around a set of themes and sub-themes initially elaborated and defined in the analytical framework. The quantitative method was applied concurrently with the qualitative method and used basic descriptive statistics on themes where quantitative data were deemed to be consistently available and comparable across the evaluation reports (i.e., QoS outputs, primarily scientific publications).

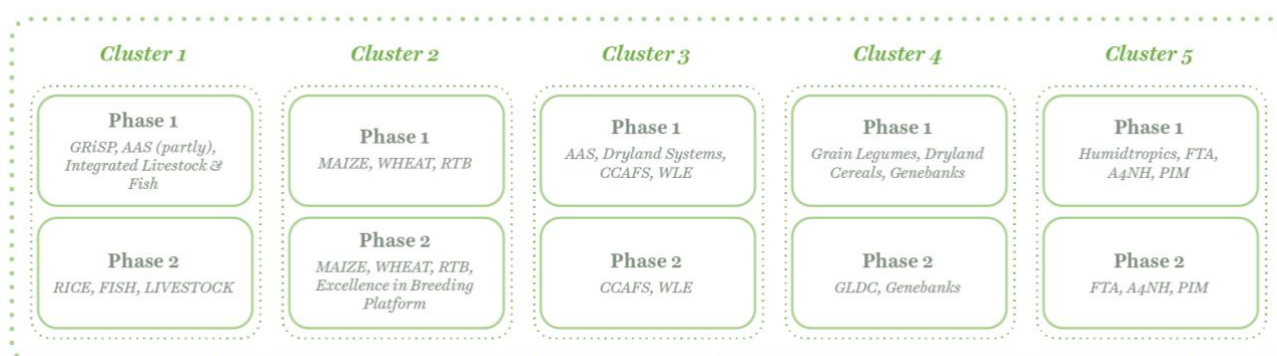
¹³ J. Popay, H. Roberts, A. Sowden, M. Petticrew, L. Arai, M. Rodgers & N. Britten (2006). Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. A Product from the ESRC Methods Programme. <https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersion1-April2006.pdf>.

The synthesis of evidence relied on an examination of 43 purposefully selected documents, including evaluations, reviews, assessments, and syntheses of evaluative evidence (among those listed in Annex 4). Most of these were externally commissioned by IEA and CAS. Based on the type and focus of documents at hand, documents were grouped into three categories, forming the basis of the three stages of coding and analysis. Five component analyses were conducted concurrently and served as the main input for this final synthesis report. The first group of 30 studies was divided into five components/clusters, according to broad groupings of similar focus and, where possible, by lead CGIAR Center. Each component group was analyzed by a subject matter expert with expertise relevant to the CRP themes concerned. This was a useful way of facilitating the description, analysis, and identification of patterns within and across CRPs, making the analysis process more manageable.

The five-component analysis (subject matter expert reports) are included as Annex 5. Most of the CRPs analyzed had two phases, but the composition and focus often changed significantly between the two phases. Component analysis results were then enriched and triangulated with information collected from thematic evaluations and assessments (e.g., cross-cutting themes, platforms)

The team used MAXQDA, a software package, for the evidence review. MAXQDA allowed the coding and analysis of specific themes as determined in the analytical framework and facilitated identifying recurring themes in the documents through the use of codes (and sub-codes) to label the themes and subthemes that emerged. It allowed for rearranging the codes hierarchically to reflect their relation to each other and editing of the names of themes as nuances were discovered during the exercise.

Figure 3. Five clusters for component analysis, by subject matter experts¹⁴



The validation of results and quality assurance relied on triangulation from different sources of evidence: the external evaluations of CRPs across the two phases, analysis of reviews of cross-cutting themes, and use of data from other reports and sources, as listed in Annex 4. In addition, recommendations were cross-referenced with those made in the evaluations and in the 2016 synthesis of phase 1 CRP evaluations. The component analyses providing raw material to this synthesis were conducted by subject matter experts, all using the same analytical framework and systematically using the coded findings from all relevant evaluations and reviews. The contributions of subject matter experts and feedback from several peer reviewers provide a deeper substantive analysis that underpins the synthesis findings and serves to strengthen the soundness of the lessons learned and recommendations articulated here.

Evidence-based recommendations were compiled by each subject matter expert, who had cross-checked recommendations by the previous evaluations and used their expert knowledge to group them; they also did their best to document how the issues identified in the first phase had been addressed, and whether they had been resolved or remained key areas for action. The recommendations arising from the five component analyses were then further validated and enriched through interactive meetings with the ISDC and CRP leaders and then brought together into this report, with their patterns and common issues identified and synthesized into the overall recommendations.

¹⁴ Executive summaries of reports by cluster from individual subject matter experts are available in Annex 5.

1.4 Limitations

Several limitations were noted and mitigated to the extent possible. Key limitations are summarized here, and more detail is available in Annex 3, Table A6. The synthesis was constrained by limited capacity and time and focused only on synthesizing existing reviews. It was therefore framed by the analyses and key evaluation questions used in sampled CRP and other evaluations. Given the heterogeneity of CRPs in phase 1 and the significant changes in CRPs in phase 2, direct comparability was not always possible (see Figure 2). Consequently, CRP evaluations and reviews themselves differed in scope and scale between the two rounds.

The variance in the Strategic Results Frameworks (SRFs) used between CRP phases 1 and 2, as well as the variance in CRP-level theories of change (ToCs) within and between the two CRP phases, constituted a challenge that the synthesis team had to handle. Developing a theory of how an intervention works, why, and for whom is the recommended approach to ground narrative syntheses. Here, the team had to rely on the evidence generated, based on the theories of change addressed in the previous evaluations and the data on uptake pathways, as available. Inferences made have been contextualized within the relevant frameworks and theories, focusing on tracing the patterns, trends, gaps, and impeding and enabling factors as well as lessons learned.

The synthesis found weak evidence in some cases and an absence of records in others (e.g., digital innovations, youth), impeding the analysis of trends and patterns for those topics. Nonetheless, an absence of evidence was not automatically judged as evidence that a particular issue was absent or insignificant. A list of thematic and methodological evidence gaps is presented in section 5.

Lastly, the synthesis evaluation did not assess the quality of evaluation reports and their terms of reference (TORs), and thus the level of potential bias could not be assessed. Additional detail on the methodology and limitations is available in Annex 3.

2 Key Findings: Trends and Lessons across the Two Phases of CRPs

This section reviews key trends and lessons from the analysis of CRPs (including platforms) listed in Figure 1. As discussed above, the main inputs for the synthetic analyses were the 47 evaluations and assessments of CRPs conducted between 2014 and 2020. Key findings from across the five components are summarized below.

2.1 Quality of Science, Quality of Research for Development (QoR4D), Inputs, Processes, and Delivery of Outputs

Following an IEA framework, quality of science was evaluated through assessment of the quality of research inputs (staff, infrastructure) and outputs (publications and other types of outputs), the appropriateness of research management processes. The quality of research for development was evaluated based on assessments of two elements: scientific credibility (robustness of research findings and sound sources of knowledge) and legitimacy (the research process is fair, ethical, and perceived as such), within the QoR4D framework (ISDC 2020¹⁵). In the synthetic analyses, special attention was given to the evolution of QoS and QoR4D between the two phases.

2.1.1 Research Inputs

Quality of research staff. Most evaluations concluded that the CRPs had highly competent, productive CGIAR research leaders and staff across both phases. However, both Dryland Cereals and Grain Legumes CRPs faced major problems in recruiting research staff for work in conflict-affected countries.

¹⁵ ISDC (2020) Quality of Research for Development in the CGIAR Context, <https://cas.cgiar.org/isdc/QoR4D>.

A number of research scientists in CRPs are internationally recognized leaders in their fields, with strong research publications records: results of the bibliometric analysis (Annex 6) show that a CRP’s highest-rated publications were often attributed to a few leading scientists.

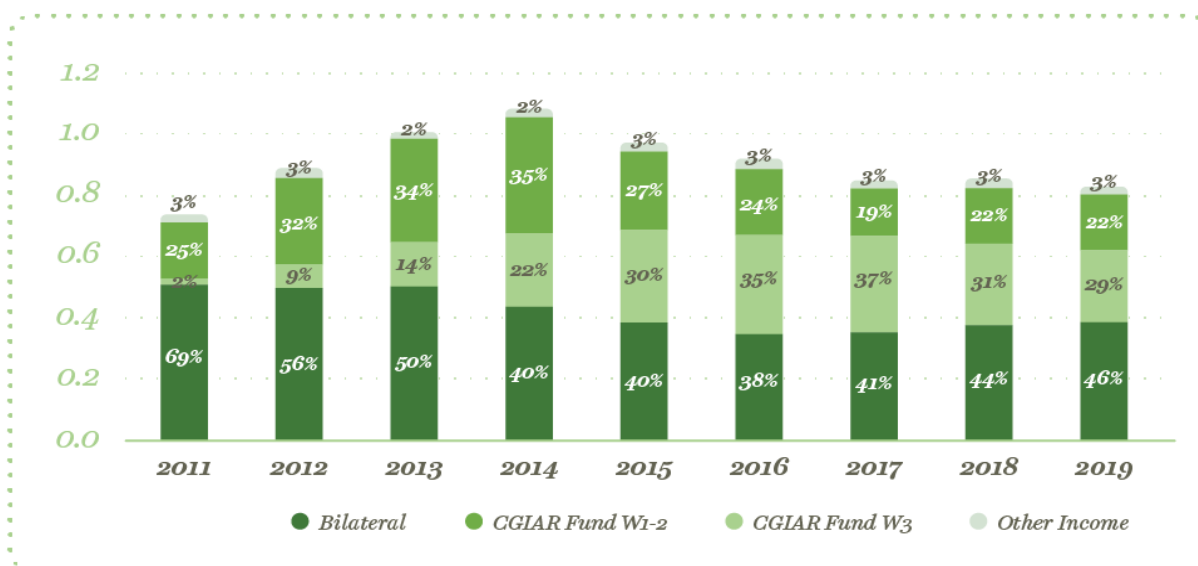
Skill mix, diversity, and allocation of staff. During phase 1, evaluations found that some CRPs had too few core scientists or lacked the range of expertise required (Livestock and Fish, FTA, A4NH, PIM). CRPs addressed many of these disciplinary gaps in phase 2 either by expanding CGIAR capacity or by leveraging expertise through external partnerships with advanced research institutions (ARIs). By phase 2, for example, the proportion of non-economist social scientists in PIM had increased to 20–30% of Flagship Program (FP) teams. A4NH gained health expertise through a new partnership with the London School of Hygiene and Tropical Medicine.

Gender balance was inconsistent across CRPs and FPs but generally improved between the two phases. Gender balance was unequal at the senior levels for RICE and Livestock in both phases but less so at the technical staff level; gender was balanced across all FTA managing partners (2nd phase). Within CRPs, some FPs had a greater gender disparity than others; e.g., A4NH’s Gender, Equity, and Empowerment Unit found it challenging to increase the proportion of male staff, despite its efforts.

Financial resources. All evaluations noted that CRPs were negatively affected by W1/W2 resource cuts, delays, and uncertainties. By the 2020 reviews, W1/W2 funding accounted for only 22% of total CRP budgets (Figure 4). In their 2016 synthesis of 15 CRP evaluations, Birner and Byerlee suggested that a minimum of 30–35% is required *“if W1/W2 funds are to provide sufficient leverage to implement an integrative and collaborative research program across Centers.”*

The W1/W2 funds were used primarily to cover gaps in high-priority areas, promote CRP coherence, and leverage bilateral funding. Resource gaps hit capacity development investments and long-term, place-based, multidisciplinary research, such as FTA’s Sentinel Landscapes, Drylands, and Humidtropics, especially hard. Between FTA’s two phases, the overall percentage of W1/W2 funds in the overall budget shrank from 39 to 10%. In the case of Humidtropics, funding reductions and delays forced local and regional partners to advance funding for research activities over long periods of time, undermining relationships and potential scaling pathways.

Figure 4. CGIAR revenue, 2011–19 (US\$ billion)¹⁶



¹⁶ Source: CGIAR financial dashboard website <https://www.cgiar.org/food-security-impact/finance-reports/dashboard/overview/>

Even as it dwindled, W1/W2 funding played a critical role in retaining a stable core of research staff (A4NH, PIM) and was also used to develop new areas of research—e.g., on food systems and human health in phase 2 of A4NH, and new scientific concepts, tools, and innovations in FTA. Across all CRPs, evaluators noted the toll taken on managers and staff by the uncertainty of W1/W2 funding and the heavy administrative load of revising budgets, seeking alternative funding, and responding to multiple uncoordinated reporting requirements.

Legacy research, partnerships, and infrastructure contributed significantly to successful CRP performance. In addition to the legacy commodity programs (MAIZE, WHEAT, GRISP), both A4NH and PIM incorporated long-running successful programs as FPs, and IFPRI's established gender research program served as the foundation for gender work in both CRPs. FISH's successful *Tilapia* genetic enhancement program dates back to 1987. CCAFS and WLE began as Global Challenge Programs, precursors of the CRPs.

Much of the research infrastructure used by FTA and Humidtropics can be traced to legacy programs, partnerships, and host-country agreements signed many years ago. GRISP and RICE research also benefited from additional high-quality research facilities through collaboration with external partners. On the other hand, FISH research was adversely affected by the inadequacy of infrastructure under the direct control and ownership of its host center, WorldFish—e.g., the molecular facilities that FISH has access to are uncompetitive because of the rapid advances in molecular technology.

2.1.2 Research Processes and Delivery of Research Outputs

The output of high-impact, peer-reviewed scientific publications was high for most CRPs across both phases. In phase 2 there was a notable increase in citations and a greater proportion of publications in high-impact-factor journals. Partnerships with ARIs, enabled through the CRP approach, were credited with enhancing the numbers and quality of publications produced, but few collaborators from national agricultural research systems (NARSs) were listed as co-authors.

The number of peer-reviewed publications per researcher was relatively low for several systems CRPs (AAS, Drylands, Humidtropics). However, Humidtropics and WLE evaluators noted the inadequacy of bibliometric analysis for assessing the quality of R4D outputs. Both produced publications about innovations and methods that had development implications but did not count toward peer-reviewed, ISI publications. While peer-reviewed publications are critical for sharing research findings within the scientific community, more accessible outputs and public media outreach may be more important for communicating with next users, implementation partners, and their networks (e.g., FTA evaluation 2014, PIM 2020, CCAFS 2020.) Most CRPs increased their output of user-accessible materials and their use of platforms such as Twitter, blogs, and Facebook to share research outputs with end users during phase 2. In general, it was challenging for evaluators to quantify the quality and impact of non-peer-reviewed communication products.

Greater integration between the CGIAR Centers and partnerships with ARIs, NARSs, and the private sector are also helping CRPs modernize their breeding programs (with some exceptions, such as FISH) through access to and use of molecular genomic and digital tools by the Centers and their partners. These advances are now being reinforced through the Excellence in Breeding Platform. Germplasm research databases have also become more uniform and accessible.

Improved germplasm of varieties/strains, parental lines, and hybrids was a key output across CRPs, with a high number of varieties released to NARSs. Over a three-year period in phase 2, GLDC produced 99 improved varieties of nutritious grain legumes and dryland cereals adapted to semi-arid and sub-humid dryland ecologies; MAIZE released 20 varieties per year; and RTB released 87 varieties over three years. More-resilient rice varieties adapted to climate extremes have also been released. The Livestock CRP achieved significant progress in both animal breeding and development of new vaccines, although success in both frequently requires longer time periods than the CRP project cycle allows.

Significant refinement of target market prioritization and product profile definition has been noted. In addition to their productive potential, a strong CGIAR legacy now carried through into the CRPs, the research programs have significantly strengthened the CGIAR System's focus on tolerance to major stresses, potential adaptation to climate change, and nutritional quality. MAIZE, WHEAT, and RICE have excelled in research toward sustainable intensification, while Livestock, FISH, and GLDC have produced and released outstanding new germplasm, such as new strains of *Tilapia* that grow much faster

than commonly used strains. The most productive programs build on key legacy programs, including the established commodity programs, such as FISH's *Tilapia* genetic enhancement program and A4NH's FP2, whose work on biofortified varieties builds directly on the long-standing Harvest-Plus program.

Despite such progress, challenges remain. Siloed approaches persisted between breeding and other disciplines and between FPs within a CRP. There was little evidence of cross-linkages between CRPs, or their coordination within a country, with some promising exceptions. PIM FP1's work with commodity-focused CRPs strengthened the biophysical parameters of its foresight models, and in turn helped commodity CRPs use the socioeconomic models in their own work. In phase 2, A4NH's new Country Coordination Units improved the coordination of work across flagships at the country level.

The GENE BANK platform has operated very differently from other CRPs. It has focused its research primarily on ex situ seed GENE BANK conservation and on developing and implementing a gene bank quality management system across the CGIAR, with associated action plans to refurbish and upgrade GENE BANKS and their operation. While these were shown to have increased accession use from the GENE BANKS, evaluations criticized the transparency of these outputs as well as the extent to which GENE BANKS reflect the genetic diversity in nature and how identified gaps will be filled.

2.1.3 Quality of Research for Development

Credibility. Considering scientific publications as an output towards measuring scientific credibility, the bibliometric analysis (see Box 1) shows that most CRPs have a strong record of high-quality, peer-reviewed publications that continued to improve across the two phases and tended to shift toward high-quality applied research. Another measure is the presence of ex ante and ex post science quality assurance processes, including structured checks on research protocols and research data management. These science quality processes are normally the responsibility of Centers, not CRPs. CRPs whose host centers had strong research management, such as IFPRI, helped to raise standards of other Centers, but the quality processes were inconsistently applied across CRPs. Some CRPs devised workarounds to solve this problem. For example, in phase 2 A4NH strengthened the quality of science by developing specific contracts outlining managing partner institution responsibilities and annual audits of performance.

Legitimacy. Ethics policies, a key measure of legitimacy, are similarly the domain of Centers and were not enforced consistently across CRPs. Progress on remedying this in phase 2 was noted. A4NH included ethics responsibilities in the contracts with managing partners. PIM required all W1/W2-funded research involving human subjects to obtain institutional review board (IRB) clearance. However, an ethical question raised in the WLE 2020 evaluation must be applicable for many CRPs: *"This reliance on real-life experiments present an ethical dilemma to many programs, with questions being raised about the fairness of using people's fields for long-term experiments, as trials are ended when funding runs out."*

There was little systematic assessment of additional aspects of legitimacy across the CRP evaluations, particularly in phase 1, although partnerships received more attention in phase 2. Several evaluations (MAIZE, GLDC) highlighted the insufficient focus on needs of the poor and marginalized in CRP research. A4NH and PIM evaluators noted the limited social analysis and disaggregated data that constrained analysis of equity and distributional issues beyond gender.

The Drylands review stated: *"Socio-economic components need to be strengthened, with poverty and livelihood assessments, adoption studies, policy and institutional analyses, and in-depth gender and youth studies."* It recommended more attention to non-agricultural livelihoods, rural-urban linkages, food systems, and policy. Similarly, the CCAFS 2016 evaluation noted that research might consider non-agricultural activities that support livelihoods and questioned whether CCAFS's implementation of the climate-smart agriculture (CSA) conceptual base was broad enough to support this.

Legitimacy of Partnerships. Perceived legitimacy is critical to the effectiveness of external partnerships. CRPs used different approaches to improve the relevance of research to next users. PIM (2020 evaluation) emphasized active collaboration throughout the research process to help ensure that its research outputs met the needs of next users. Several CRPs prioritized close links and responsive partnerships with governments at the policy level (MAIZE, WHEAT, PIM, A4NH, FTA). Other CRPs focused including downstream players and the resource-poor. AAS, WLE, Drylands, and Humidtropics all featured close engagement and fieldwork with communities, in the case of the systems programs centering on specific landscapes and communities, while Humidtropics further emphasized multi-stakeholder processes

to improve the relevance of social and technical innovations for specific agroecological systems. However, the efforts of the systems CRPs were short-lived and none was renewed for a second phase.

Coordination with country partners and alignment with national policies and strategies remained challenging for most CRPs. Dialogues were facilitated through the Global Conference on Agricultural Research for Development (GCARD) at the start of phase 2, enabling collective national dialogues in more than 20 countries between the relevant CRPs and their national partners. Despite positive initial intentions to coordinate with country policies and systems and to work through a unified approach and avoid duplication, there was little evidence in the reviews that these discussions resulted in practical steps such as common focal regions for research or unified planning, resourcing, and capacity development within a country.

Box 4. Bibliometric analysis

Bibliometric analysis is a key proxy measure for evaluating scientific credibility. It was undertaken to assess the quality of scientific publications produced by CRPs over the two phases. A key limitation of the analysis is the varying comparability and availability of data for the two periods. The conclusions of this analysis are summarized here; detailed results are provided in Annex 6.

The analysis focused on four core bibliometric indicators, selected according to relevance and considering the limitations above.

Citation analysis: The citation analysis categorizes all CRP-specific publications based on the number of times they were cited during the respective analysis period, across eight brackets. Aggregated, the data show an increase in the overall number of publications between the two phases. However, when comparing the normalized distribution across citation brackets between 2015 and 2020, the increase in publications is disproportionately distributed. Notably, the percentage of articles with 0 citations decreased by one-third, from 25.1% to 16.1%. Articles in the 1–10 and 11–20 citation brackets increased by 13.9% and 13%, respectively. Among the more cited brackets, 21–30 and 31–40 showed decreases of 27.7% and 6%, respectively. Finally, the brackets with the most highly cited papers (i.e., 41–50, 51–99 and >100) increased by 54.8%, 74.2%, and 100%. The latter data underscore the increase in highly cited publications since 2015.

Most-cited articles: Most-cited publications were examined for every CRP. Over the period 2015 to 2019, the average number of citations of the CRPs' most-cited publications increased by 11%.

Journal frequency: This metric is the number of publications in the three most frequently published journals for each CRP, including the impact factor (IF). The number of publications increased by 3% between 2015 and 2020, while the average IF of the three most frequent journals increased in two out of three cases, where data are available. In 2015, the most frequent journal was *Theoretical and Applied Genetics*, with 43 articles (14.6% of publications in the three most-published journals). In 2020 the most frequent journal was *Field Crops Research*, with 45 articles in 2020 (14.8% of publications in the three most-published journals).

H Index: This metric is based on a subset of researchers at each CRP, including the number of people and the highest H Index within the selected group. The comparison between 2015 and 2020 is difficult because (1) the subset of researchers analyzed for each CRP varies between CRPs and between the two time points, and (2) the 2015 data were obtained from Scopus, while the 2020 data- from Web of Science. If only the maximum H Index is considered, there were significant increases for all CRPs where data are available.

Generally, across the CRPs, while many positive examples of benefits from CRPs' integrated approaches to improving food and nutrition security were found, evaluations concluded that research was largely not focused on the needs of the poorest and most vulnerable sectors (e.g., MAIZE, GL and DC) and often gave little prominence to potentially valuable *off-farm* income and employment opportunities for rural women and youth around agri-food systems (although youth enterprise platforms have begun to appear). This gap also reflects limited social science capabilities in the CRPs and a predominant focus on biophysical dimensions rather than on alleviating poverty and improving the lives of the most disadvantaged (SLOs).

2.2 Performance: Delivery to Outcomes, Objectives, and Sustainability

Overall, the CRP approach has brought more coherent planning that has helped to increase the development impact of CGIAR research. The TOC and flagship programs (Annex 2) have focused actions toward specific, readily recognized outcome targets and leveraged partner and funding commitments accordingly. However, where individual CRPs and FPs were not well resourced, there were significant gaps in the ability to deliver coherent outcomes, as seen for example in Livestock FP3, GLDC FP2, FTA's Sentinel Landscapes, and Humidtropics.

2.2.1 Delivery of Outputs and Outcomes

CRPs delivered an array of outputs and outcomes at different levels. Only in a few cases were there joint actions between CRPs. Most of the evaluation methods focused primarily on scientific inputs, activities, and outputs.

Commodity and systems CRPs. Stepwise research milestones were well documented and generally achieved. Commodity-focused CRPs were found to apply high levels of scientific inputs into accelerating breeding cycles and achieving higher genetic gains through genetic interventions. Increasingly, work on sustainable intensification has involved and optimized Genotype x Environment x Management for varieties/strains under a wide range of conditions essential for local adaptation and sustainability. Closing the yield gap through sustainable intensification is critical to future improvements in farm productivity and profitability, and major scientific input to that area is evident in the MAIZE and WHEAT CRPs. The relatively lower dedication to understanding and promoting sustainable intensification of cropping systems in RTB may constrain pathways to reducing yield gaps in target regions. For rice, climate change has become a key focus; work includes breeding for resilience to stress conditions and agronomic measures to reduce water consumption and carbon release from lowland systems.

The CRPs that focused on system-level R4D approaches at the field level (e.g., participatory research, action research) faced the most severe difficulties in delivering the outputs and outcomes they promised. According to the evaluations, these challenges stemmed, first, from unrealistic planning and expectations. For example, Humidtropics' evaluators noted the solid theoretical underpinnings for its systems research and found that it produced high-quality outputs (manuals, tools) to facilitate stakeholder engagement, but it was slow to test technical innovations at the system level, and the end of phase 1 was too early to assess progress toward outcomes. Second, the challenges stemmed from design flaws. For example, in implementation AAS had limited emphasis on innovations that enhance systems productivity or on poor and marginalized people, affecting its potential for development impact. Drylands systems research activities were criticized for focusing on incremental improvements in existing farming systems instead of game-changing system innovations. Several lessons for future system transformation work in the CGIAR emerge: sufficient time must be allowed; feasible targets should be created; and the design must be framed in terms of a pro-poor and transformational perspective from the start.

Policy outputs and outcomes. By phase 2, evaluations indicated good progress on policy outputs and outcomes at different levels, particularly from the global integrating programs. Research from these CRPs has contributed significantly to policy discussion and change related to climate change, nutrition and health, response to economic shocks, and the sustainable management of land, water, and ecosystems. By phase 2, A4NH had developed many innovations and policy contributions, and gender tools and other social science outputs. FTA supported the development of global policies on agroecology, adaptation, and forest genetic resources. CCAFS and WLE phase 2 reviewers commended the achievement of planned outputs and outcomes and noted the CRPs' respective roles in catalyzing climate change action and contributing effectively to the sustainable management of land, water, and ecosystems.

Are assumptions about targets, timeframes and CRP contributions reasonable? FTA's phase 1 evaluation noted that while FTA projects usually delivered outcomes that were within its direct control, performance on larger-scale development outcomes was less than satisfactory *"given the assumption that pilot-scale achievements would, by themselves, become effective drivers of replication, adoption, and further applied research."* The phase 2 evaluation reported improved quality of FTA partnerships but lacking evidence of large-scale development impact, largely because of missing scaling-out partnerships and resources.

The 2020 PIM review noted the CRP's significant progress on three outcomes common to the flagship programs: capacity development of research partners and poor and vulnerable groups; contributions to more gender-equitable control of assets and resources; and informing of policymaking at different levels. The 2015 evaluation of PIM cited the lack of support for research on the science-policy interface, which was addressed through new research products including the Kaleidoscope Model and accompanying case studies. However, the review noted the continuing challenge of capturing the extent of PIM's contribution to policy outcomes, especially distinguishing the importance of research outputs versus non-research influences on policy change. Logically, other CRPs that focus on contributions to global policy discussions, documents, and outcomes—CCAFS, WLE, A4NH, and FTA—will all face this same challenge.

The case of A4NH illustrates the challenges around setting and meeting achievable targets that correspond to SLOs. In phase 2 A4NH set three targets for its contribution to the second SLO, improved food and nutrition security for health. Significant progress was made on two targets, largely attributable to FP2, Biofortification. For Target 1, by 2019, 8.6 million households had adopted improved varieties or management practices (compared with a target of 20 million households by 2022); 8.5 million of these were households growing biofortified crops. For Target 2, 42.4 million people had benefited from biofortified crops by 2019 (compared with a target of 150 million people without deficiencies of one of more targeted micronutrients). For Target 3, there was no evidence by 2019 of progress toward improving diet diversity, against the target of 10% fewer women of reproductive age across four target countries consuming less than the adequate number of food groups.

Across CRPs and FPs, progress on outputs, outcomes, and scaling usually reflected the different maturity levels among the programs. The most advanced programs (and outputs/outcomes) were established many years before the CRPs. Solid partnerships and legacies generated through both the reputation of the Centers and the work of individual scientists with long-standing partners have brought trust capital and ready mechanisms to facilitate the achievement of outputs and outcomes.

A straight progress indicator against targets does not fairly account for the time needed to establish new programs, such as the systems-level CRPs or FP1 in A4NH on dietary diversity. Careful assessment is needed to determine whether long-term targets are realistic, measurable, and achievable within a given timeframe and with CRP capacity.

A related challenge in assessment arises when targets include a heavy delivery component involving other actors. What is the appropriate role of the CRP and CGIAR vis-à-vis other partners external to the CRP? How can the CGIAR's role be defined more concretely and appropriate indicators be constructed to assess progress made within this role? How do partners and other stakeholders perceive and evaluate the success of the CRP?

2.2.2 Limited Coherence in the Use of ToC

For almost all CRPs, the reviews showed significant development of technologies and innovations for potential adoption. In general, however, evaluations found it difficult to determine the linkages between FP objectives, outcomes, and the CGIAR-level SLOs they are intended to contribute to. TOCs were used in the initial planning of phase 2 research but little used to guide the management of programs to deliver strategically toward defined sustainable development outcomes. Among the CRPs reviewed, only FISH explicitly determined success in terms of achievements along the TOC pathways to SLOs. For example, an additional 400,000 metric tons of fish was produced, decreasing water use by 37% and cutting greenhouse gas emissions by 22% (SLO 3.1 target was an increase of 4.8 million metric tons; FISH met 8% of the target) (2020 evaluation of FISH).

The other CRPs used milestones to track their achievements that were not well linked to their ToCs and the CGIAR SLOs. Milestones (phase 2 only) were set at a range of levels, from activity to impact and thus were not useful for indicating progress toward IDOs/SLOs. Similarly, outcome impact case reports (OICRs) (also at phase 2 only) have been helpful to highlight key achievements, but by themselves they cannot provide a comprehensive and accurate gauge of progress from outcomes to impacts.

The question is not whether the CRPs have contributed to outcomes—they have. All the CRPs have some examples of measuring pathways toward impacts to varying degrees of robustness. But the evidence provided does not show the extent to which the results have contributed to the sub-IDOs, IDOs, and SLOs specified in their ToCs, by which they secured funding.

Why were the TOCs not used more appropriately or effectively? The evaluations point to a range of factors: (1) quality of TOCs varies among FPs, and they are too often generic; the 2020 reviews reported flaws in the CRP TOCs and stated that the reporting system is weak, with unnecessary duplications and inadequate means for assessing the CRP's contributions against the TOC; (2) there was a lack of technical support for developing and using the TOCs to guide monitoring and evaluation; (3) there were pressures to develop more user-friendly products to communicate progress and achievements with partners and funders, (4) research managers facing multiple, uncoordinated sets of reporting requirements suffered from time constraints; (5) key stakeholders lacked ownership, and top-down approaches prevailed; (6) not enough attention was given to testing the assumptions and risks identified in the ToCs as dynamic concepts to be checked continuously; and (7) ToC narratives do not usually establish links with existing literature, which lower their credibility in the eyes of academic community.

A key lesson emerging from the evaluations is that the likelihood of outputs being scaled out/up increases if the CRP has:

- A comprehensive understanding of the multifaceted needs of targeted end users in a changing environment (e.g., GRIISP's sound priority setting sharpened its focus on appropriate stress factors while maintaining a productivity focus)
- Informed support of scaling-up activities and strong links with next-stage users (e.g., in LIVESTOCK, cattle-feeding research brought together investment in the dairy sector by the Government of Kenya and USAID to jump-start development of a whole subsector in Kenya)
- Developed outputs that are readily transferable as being relevant across countries with similar agroecological and socioeconomic conditions (e.g., high-yielding rice varieties can be readily transferred between Asian countries with the same "seeds without borders" network agreement but are not suited to transfer between continents owing to differences in biotic and abiotic stresses and consumer grain quality preferences).

2.2.3 Insufficient Research, Measures, and Skills for Sustainability and Scaling

Inadequate measures. The effectiveness evidence provided for most CRPs is in the form of impact assessment studies and OICRs, which, as discussed above, do not explicitly align with, or show the contribution to, the IDOs or SLOs specified in their ToCs. In some cases (e.g., MAIZE), impact studies lacked adequate quality and credibility. Weaknesses in monitoring, evaluation, and learning (MEL) systems were apparent in many cases, with measures frequently confined to release of technologies rather than following through to assess their take-up by resource-poor farmers or their use in policy changes.

Current CGIAR metrics for development impact, sustainability, and resilience are insufficient. For natural resource management (NRM) and systems research, progress indicators and impact assessment methodologies are challenging, and it takes significant time to achieve impacts, often exceeding a typical project or CRP duration. Evaluations of WLE in 2016 and 2020 noted that "efforts to assess the impact of NRM research in the CGIAR are recent compared to the considerable work and methodologies developed for assessing the impact of genetic improvement."

The lack of appropriate MEL for the development objectives of CGIAR is a significant handicap in seeking to demonstrate the extent to which outcomes are attributed to or contributed to by the CRPs. Metrics on resilience, poverty alleviation, and sustainability have been still less used, reflecting the lack of social science and of measures that are easy to apply. Evaluations and reviews express a clear need for a more streamlined and appropriate approach to measuring progress and learning.

Box 5. Scaling in the CRPs

Scaling, as used in this report, is defined as “expanding, adapting and sustaining successful interventions (policies, processes, programs or projects) in different places and over time to reach a greater number of people.” The CRP assessments found that the likelihood that outputs will be scaled out increases if the CRP has:

A comprehensive understanding of the multifaceted needs of targeted end users in a changing environment – for example, GRiSP (building on legacy work) focused on stress factors while maintaining productivity traits, resulting in submergence-tolerant rice varieties which reduce yield losses in flood-prone areas.

Informed support of scaling up activities and strong links with next stage users – The Livestock CRP’s research on improved cattle feeding practices to increase milk productivity in Kenya was adopted by 80,000 farmers. This work built on 6 years of previous programmatic work and investment by USAID and the Government of Kenya in better genetics for dairy animals, improved health management, new breeds of forage grass, clearly defined and successful business models, and strong collaboration with government and nongovernment partners. It provides an important example of how sets of innovations and technologies can align with developmental investment and capacity building to facilitate change at scale.¹

Developed outputs that are readily transferable – GRiSP experiences demonstrate the importance of local adaptation and using a landscape lens. GRiSP increased its efforts to develop sustainable rice management technologies such as direct seeded rice, alternate wetting and drying, and laser levelling. However, the site-specific nature of many practices has inhibited widespread adoption. Much of the research focused at single field level, with limited attention given to the consequences at landscape level. Although AWD has clear water-saving benefits, especially when applied at a landscape level, national partners in Cambodia, India, Thailand and Vietnam stated that coordinating the practice among neighboring farmers was difficult. Similarly, replication of widespread uptake of high yielding rice varieties across different agro-ecological zones is problematic. High-yielding rice varieties developed for, and readily adopted in, Asia or Latin America are not necessarily suitable for Africa because of differences in biotic and abiotic stresses and consumer grain quality preferences.

Research and skills to support sustainability and scaling. CCAFS’s 2020 evaluation concluded that most innovative emerging strategic research concerns the question “What processes can facilitate adaptation, resilience, and mitigation?”— i.e., the science of delivery or social learning—as opposed to “What techniques can be developed and deployed to underpin adaptation and mitigation?”

WLE and systems CRPs, including Humidtropics, emphasized the importance of convening processes “that can bring science and development together, in particular engaging with diverse stakeholders to build mutual understanding and agreement on the potential of different options to manage landscapes and improve agriculture” (WLE evaluation 2016). The skill set and type of engagement needed to move from outputs to impact at scale is very different from the training and background of most CGIAR scientists.

A frequently noted bottleneck constraining the contribution of research outputs to sustainable development has been the lack of capacity in national agricultural research and extension systems (NARESs). Individual capacity development for national scientists has featured strongly in all the CRPs, mostly through bilateral funding support, as has training of farmer groups. The Excellence in Breeding (EiB) platform is recognized as having helped NARESs incorporate genomic tools in their work in RTB, in addition to enhancing capacities in the CGIAR Centers. Nonetheless, it was recognized in all evaluations that capacity development in uptake partners remains a critical need, requiring more strategic and systematic focus on developing research and innovation capacities at all levels and fostering institutional change to address the innovation webs involved. Unfortunately, coherent capacity development strategies at the institutional level, not only for individuals, and strategies to hand research processes over to national systems were found to be largely lacking in the CRPs. External capacity development was one of the first areas cut when funds were reduced. Findings on capacity development show the need for more equitable relationships with partners, so that all “own” and can themselves commit to the agenda and processes involved.

Humidtropics' research on stakeholder processes and innovation platforms is an example of the type of research that can contribute to successful systems adaptation and transformation, and to scaling and sustainability processes that are 'owned' by stakeholders. The research featured a dynamic process of local appraisal of the community by existing and potential partners, work with them to identify entry points for change, and the development of working partner alliances. The approach enabled stakeholders to share ownership of the research process while facilitating community empowerment as the range of partners expands (HT evaluation 2016).

Evaluators also noted the importance of partnership for policy innovation and impact. For example, the ability of PIM and A4NH to contribute effectively to policy processes at national levels was heavily linked to longstanding in-country partnerships and relationships of trust developed over the years by the host center IFPRI. Partnerships are also essential for the development of policies that support research, for example to enable the international exchange of germplasm. Here, insufficient engagement was noted at both the national level, with country policies restricting the international transfer of germplasm to benefit others, and at the global level, by the GENE BANK platform and CGIAR System, with global treaties such as the International Treaty on Plant genetic Resources for Food and Agriculture (ITPGRFA), the UN Commission on Genetic Resources for Food and Agriculture (CGRFA), and the Convention on Biological Diversity (CBD). Partnership were enhanced through the platforms, which reinforce the collaborative nature of the CGIAR's work, both between Centers and with external partners. However, a 2017 evaluation found the GENE BANK platform was not meeting needs of NARS partners; support is maintained only for CGIAR's ex situ collections, with the evaluation gap as to how it addresses in situ conservation of crop wild relative and landrace populations. The ownership and use of banked germplasm remain sensitive issues for national systems and need addressing in future actions and evaluations.

Engaging the private sector. The CRPs have focused primarily on government and community partners, and partnerships with ARIs, but engaging the private sector will be critical for scaling innovations. Work with the private sector expanded during phase 2 across a number of CRPs, although there was no evidence of a private sector engagement strategy or analysis of the effectiveness of these efforts and lessons learned for scaling and sustainability.

Expansion of work with national seed systems during phase 2 was particularly promising. During phase 2, RICE FP1 and FP2 strengthened seed systems and delivery efforts in response to gaps identified during the CRP design process. RICE is now working with 20 partners—half NARES and half private sector companies—for seed delivery in South Asia and Africa, with similar work underway in West and Central Africa through AfricaRice (RICE 2020).

Market segmentation is an important consideration; for example, cassava adoption from the RTB CRP has been much more rapid in Vietnam, where new markets are developing for commercial starch production, than in Nigeria, where the prime market is smallholder domestic consumption. While target market prioritization was noted as having improved through the CRPs, reviews of GLDC and MAIZE noted that a focus on high productivity in high-potential systems can be effectively addressed by the commercial sector and CGIAR's focus should be more on the needs and risk awareness of resource-poor farmers. Evaluations of WHEAT, MAIZE, and RTB noted that engagement with private sector partners needs to be further enhanced to address defined market sectors and make use of private sector comparative advantages in delivery to market needs.

CRPs have clearly recognized the importance of partners in impact pathways, resulting in some innovative partnership arrangements, but these remain linear in thinking rather than cyclical processes of innovation, learning, and evidence-based feedback.

Engagement with wider development investments and programs. As discussed elsewhere, engagement with national strategies, plans, investments, and official development assistance (ODA) has been patchy and incoherent. There are good examples of CRP research objectives having been strongly owned and committed to by national systems. In MAIZE, outcomes from three case studies showed clear contributions to CRP IDOs, with direct large-scale investments by national system partners (e.g., through MasAgro) to ensure that CRP outputs addressed their specific needs and could be readily taken up. However, given that around half of all farmers are considered unlikely to grow out of subsistence farming

directly¹⁷ and need off-farm opportunities in rural areas and small towns, the relative lack of connection to rural development processes—and of an associated focus on agri-food employment and enterprise beyond production—remains a major gap in CGIAR’s pro-poor work.

2.2.4 Management and Governance

CRPs were usually highly dependent, and benefited significantly from, the lead Center’s well-established governance, finance, and human resource systems, and research programs. However, in some CRPs (e.g., PIM, GL, DC) questions arose in phase 1 over the lead Center’s dominance, specifically over transparency of funding allocation and potential conflicts of interest given the lead Center’s role in governance as well as its fiduciary responsibility. The Fund Council and Consortium agreed on new governance structures for the next generation of CRPs at the end of phase 1, with each CRP having an Independent Steering Committee as its central decision-making body. Phase 2 assessments noted the positive impact of the ISCs in improving trust and cooperation among researchers, synergies in research activities and funding transparency, and an independent source of advice. Nonetheless, in some cases, funding remained closely linked to particular Centers as sources of expertise; for example, in 2020, IFPRI received 75% of the W1 & 2 funding for PIM, but also mobilized over 90% of the W3 funding used.

There are important lessons from CCAFS, which evolved into a more independent structure, with the program management unit (PMU) outside the lead Center and expert non-CGIAR staff recruited for leadership positions. CCAFS’s unique path bolstered its identity as a semi-independent integrative program and allowed it to focus on the strategic partnerships needed to achieve program goals, bringing fresh perspectives and increasing its capacity and the quality of its science (CCAFS 2020).

Through the CRPs, cross-Center collaboration has now become established as a standard way of working. In RICE, CRP links between scientists have broken down many institutional barriers between researchers in Asia and Africa. As a result, IRRI and AfricaRice have moved into a much closer alliance, facilitated by their scientific links. Similarly, RTB, MAIZE, and WHEAT have facilitated stronger inter-regional linkages among the Centers involved. PIM brought together and strengthened the community of social science researchers throughout the CGIAR. A4NH has linked researchers and research across the Centers to elevate the CGIAR priority to nutrition.

GENEBANK has promoted synergy between Centers and CRPs and supported breeding work across flagships and regions with effective methods for sharing methods, tools, and experience. It represents a unique link between CGIAR and the Crop Trust, bringing much more stable funding than before to the maintenance and use of CGIAR germplasm collections. Nonetheless, evaluation has identified potential for Crop Trust and CGIAR conflicts of interest in (1) programmatic and financial governance, (2) policy development and representation, and (3) communications.

2.2.5 Conclusions from the Analysis of Key Findings, and Trends

The CRPs have added clear value to the CGIAR. They have, among other things:

- Increased CRP and Center collaboration, creating new research synergies between Centers and research teams that had not previously worked together
- Increased the breadth, depth, and strength of partnerships including, but not limited to, world-class ARIs, other CRPs and CGIAR Centers, national innovation and policy systems, the private sector, civil society, and farmer organizations
- Expanded relevant and innovative multidisciplinary research
- Addressed and expanded new thematic areas for CGIAR, including nutrition and health, market chains, climate change, and systems-based approaches using research in development
- Taken forward traditional areas of strength such as plant and animal breeding, with a focus now going beyond productivity to address new challenges such as climate change, water scarcity, and invasive pest outbreaks

¹⁷ DFID, DFID’s Conceptual Framework on Agriculture, 2015, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/472999/Conceptual-Framework-Agriculture2.pdf.

- Identified systemwide gaps and instituted new platforms to sustain and grow actions in these areas, including GENE BANKS, breeding technologies, gender and women's opportunities in research and on the farm, and new opportunities for use of the big data holdings of the system
- Confirmed the international profile of CGIAR and the value of agri-food research through evidence-based policy-level engagement, and expanded its relevance in areas such as climate change, natural resource management, and production of varieties and practices to foster resilience and sustainable agriculture
- Added significant value to existing CGIAR expertise, programs, and relationships through expanding research and partnerships
- Displayed flexibility and speed in responding to new challenges, such as fall army worm in maize, by mobilizing partner capabilities in and beyond the system.

Across the CRPs, evaluations have recognized that the solid performance of the CRPs in bringing together Centers, mobilizing new partners, dealing with highly variable funding and moving strategic targets, taking on new areas of research, and delivering high-quality scientific products is a real credit to the managers and staff involved. The CRPs have required whole new ways of matrix management, inter-Center collaboration, and approaches to bringing together separate research efforts into an integrated whole, driving coherently toward common outcomes and impacts.

3 CGIAR System-Wide Lessons and Conclusions

3.1 Impact Pathways of CGIAR Research

The CRPs have produced high-quality and relevant research products, but there is a disconnect between the time to impact and the lifespan of a project or CRP. From one side, the long-term and complex nature of the research-to-impact pathway make it difficult to obtain the commitment of partners required to enable uptake into use. From the other side, donor expectations put pressure on the CRPs to articulate short- to medium-term development outcomes that are unrealistic for CGIAR research by itself and put weight on the CGIAR Centers to focus on more adaptive research, for which they may not have a comparative advantage, at the expense of longer-term exploratory research.

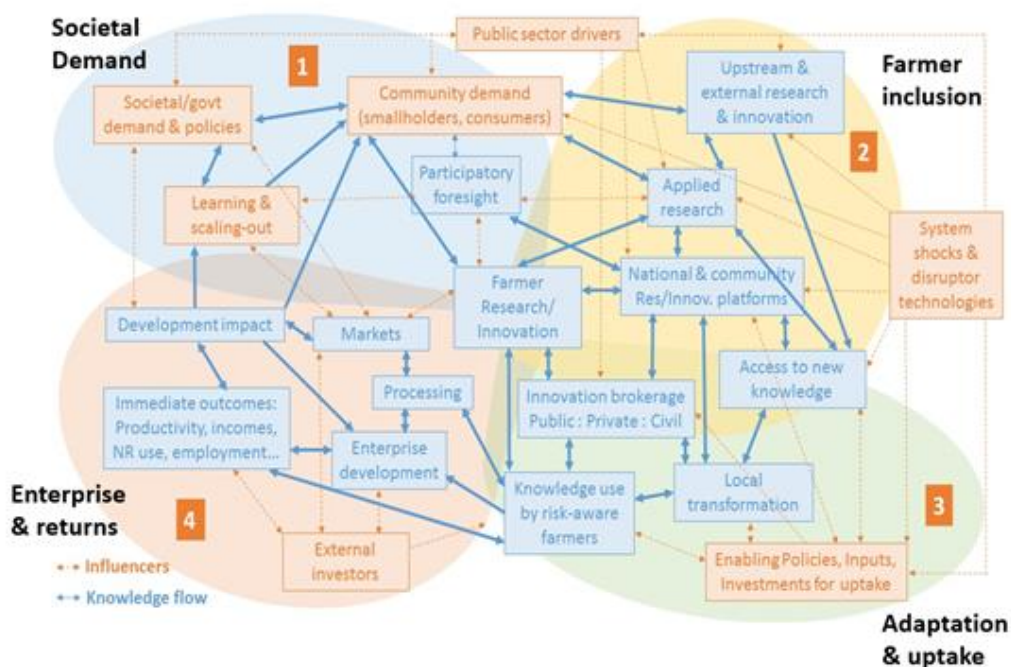
Results chain from FP objectives to CRP ToC and to broader-level CGIAR SLOs should be strengthened. Progress was stated in strengthening the rationale for and coherence among FPs in several CRPs. Nonetheless, the linkages between flagship objectives, outcomes, and broader-level CGIAR SLOs were often not evident. The CRPs reported against milestone systems and OICRs for each flagship, but these were not usually linked to ToCs. The milestones were used for planning purposes but kept changing, further compromising their use as indicators of progress along the program ToC.

Impact depends on key stakeholders' engagement along the impact pathway, national public and private partners, and intended end users. Several CRPs (e.g., Livestock, Fish, and Rice) have been assessed positively in enveloping and strengthening partnerships with key players along the impact pathway in order to achieve the CRPs' respective objectives. While there are some excellent examples of partnerships that have led to impact, a more systematic and systemic approach to engagement with stakeholders is still needed, especially in terms of reality checking and early engagement.

Likewise, impact is a function of the end users' capacity to adopt. The CRPs collectively have increased their knowledge of the needs of the end users (farmers, agri-businesses, policy makers, and so on), either through targeted research or as a result their in-country lived experiences, but given today's rapidly changing social, economic, and natural environment more could be done, especially in terms of managing risks to the farmers at the onset. Along the same lines, capacity development and partnership/network development can affect CRP progress along the impact pathway. Still, metrics are overwhelmingly focused on countable outputs such as trainee numbers and partners engaged rather than more subtle but arguably much more important outcome indicators.

With the exception of externalized innovations such as vaccines and prescribed delivery pathways such as seed supply systems, there is an innate flaw in presuming linear pathways to impact for the wider sustainable development agenda, where innovation takes place through many drivers and from many sources. It involves webs of innovation, with many actors involved in shaping priorities and in enabling the development, transformation and integrated use of innovations through the multi-directional sharing of ideas and cycles of learning and feedback (Figure 5). CGIAR provides an important entry point through innovations and learning from international science but meeting the needs of the resource-poor requires such knowledge to be embedded in the wider reality of sustainable development. With three-quarters of farmers in 'developing' countries farming less than one ha (and 85% less than two ha)¹⁸, smallholder farms fulfil multiple functions and are central to the livelihoods of rural communities. Hence also: "sustainable intensification among smallholders is not just another optimization problem for ensuring higher productivity with less environmental impact. Rather, it is a complex task of creating value through innovations in the institutional, organizational and technological systems of societies."¹⁹

Figure 5. Multiple entry points for participatory innovation



3.2 Monitoring, Evaluation, Learning, and Impacts Assessments (MELIA)

Across the CRPs, while inputs, outputs, and their scientific quality have been well measured and outcomes and impacts fairly determined (partially via SPIA) for linear uptake pathways of some new varieties and technologies attributed to CGIAR, the same is not true of CGIAR's contribution to wider and

¹⁸ HLPE. 2013. Investing in smallholder agriculture for food security. A report by the High Level Panel of Experts on food Security and Nutrition of the Committee on World Food Security, Rome

¹⁹ F.Gatzweiler & J von Braun (2016) Innovation for Marginalized Smallholder Farmers and Development: An Overview and Implications for Policy and Research 1-25 in F. Gatzweiler & J von Braun (eds) Technological and institutional innovations for marginalized smallholders in agricultural development. Springer International, pp435

more complex development outcomes, where evidence gaps were identified across a wide range of CRPs and themes.

The 2017 evaluation of results-based management (RBM) noted the need for a more compelling strategy and results framework (SRF) and the introduction of sub-intermediate development outcomes (IDOs) to elaborate on IDOs as a way of scoping expected CRP outcomes within a more reasonable sphere of influence and better connecting research and development. MEL system was recognized to have promoted accountability in meeting milestones and reporting deadlines, yet strong recommendations for better MEL in tracking of IDO/SLO delivery were made for at least half the CRPs after phase 1. Similar calls were made in 2020, when evaluations highlighted the need for more useful approaches to show delivery against the SRF, more use of qualitative analyses alongside quantitative data, comparative studies of delivery models, and determination of the outcomes of policy interventions and capacity development processes.

Reporting systems, including MARLO,²⁰ were considered important but deemed to be overly complex, very time consuming and requiring a duplication of effort, without providing a complete picture of CRP progress and lesson learnt. Besides, the CGIAR reporting system is perceived as “imposed,” is subject to frequent changes, and requires a great deal of personnel support. In the specific case of non-CGIAR centers, MARLO was attested to not fully capture their contributions (Rice, 2020).

While there was evidence of learning from outcomes by the CRP and external partners, the learning component of MELIA was not well formalized. This meant that while valuable learning did take place according to the evaluations, it was not necessarily recorded or built on.

Evidence is short on whether and how results from on-farm and adoption studies were used to feed back into CRP design and prioritization for improving relevance and effectiveness. Reasons for non-adoption of technology are not always adequately reported and translated into specific recommendations. At the same time, there is little evidence that the FPs have revisited their ToCs for planning, learning, or adaptation purposes.

The number of impact assessments carried out is limited. The main reported factor was the lack of funding. Impact evaluations by their nature are expensive and require long-term planning whereas CRPs operate on the basis of annual planning schedules. This means that various sources of funding are required. Also, the tendency to prioritize producing innovations over carrying out evaluations was another hampering factor in cases where responsibility for evaluation lay mainly with the technical FPs, which do not necessarily see it as a priority as they are “science focused.”

“Everyone is eager to use evidence, no one is willing to pay for it.”

Livestock CRP Review 2020 Report

Independence (or perception of independence) of assessments is key for accountability and results uptake. In some cases the level of independence of impact assessments was questioned, particularly when they were led by an FP, even with participation or advice from external stakeholders. While improvement of FPs’ capability to conduct robust assessments over the years was a positive result, the risk of a conflict of interest was recognized as high by evaluators.

3.3 Greater Social Science Engagement to Bring Clearer Focus on the Poor and on Scaling Processes

CRP research has been strong and of internationally recognized quality in its biophysical work but markedly less so in dealing with the “people” dimensions of development. This has resulted in insufficient attention to equity issues beyond gender on the whole, and limited capacity to grapple with the socio-cultural-political factors that affect scaling and sustainability. With limited social science involvement, there has been a continuing tendency for breeding processes to be technology-driven and operate in their own silos rather than being strongly linked with sustainable intensification processes and their wider socio-agronomic implications. Social science needs to be given a much stronger role in the formulation

²⁰ MARLO is a management tool that aims to facilitate both the monitoring of progress toward the SRF SLOs and the production of CRPs’ operational documents (POWB and Annual Report).

and delivery of research for development programs. The CRPs have struggled to recruit internally in this area, but several (A4NH, PIM, CCAFS) have been able to fill disciplinary gaps through innovative multiyear relationships with external organizations.

3.4 Ownership of the Agenda by External Partners and Stakeholders

The case is strong for participatory foresight and planning in determining research priorities and ensuring ownership of preferred future scenarios by all concerned parties, so that all can commit, innovate, and work together toward these ends. Rather than a research-centred agenda based on linear assumptions of uptake, and prone to successive barriers to success, more equitable and inclusive processes of prioritization and partnership recognize the complex reality of development and the innovation webs involved in scaling, transformation, uptake, and use and starting from development agenda and national commitments in which research products can be embedded and used.

3.5 Resourcing and Measuring Contributions of the CGIAR and Partners to Development

The separation of CGIAR central funding from donor bilateral commitments to countries works against coherent linkage to development programs and needs to be addressed within donor agencies themselves. Reorienting CGIAR support to align with development commitments can be very effective in ensuring the value of outputs is realized in sustainable development (as seen in MasAgro in Mexico). From validation discussions, the CRP leaders clearly recognize this, but a policy-level shift in practice and evaluation is required in CGIAR from seeking to **attribute** development impacts to a Center's research, toward determining and valuing the essential **contribution** CGIAR is making with others through its research.

4 Actionable Recommendations

4.1 Recommendations for One CGIAR

These recommendations build on evidence-based conclusions by the subject-matter experts²¹. Many map directly to recommendations from previous evaluations, thus still remain relevant.

- 1** System funders and One CGIAR managers should invest in preserving and taking forward valued elements developed through the CRPs: infrastructure, relationships, processes, tools and innovations.
- 2** Prioritize partnership development and stakeholder engagement. Develop and implement a system-wide strategy for equitable engagement and effective communication with partners and stakeholders of all categories in the foresight, planning, delivery and follow-through of CGIAR research, with metrics derived from partner perspectives.
- 3** Focus much more on institutional capacity development, especially of national 'boundary' partners, in close collaboration with donor agencies and other funding partners. Develop and implement a system-wide strategy and partnerships with other agencies to facilitate development of required capacities for uptake, transformation and use of CGIAR products, through capacity development. To help achieve development outcomes, CGIAR and its programs should more actively advocate and help leverage financial resources for capacity development of national partners in pathways to impact. This requires a clearer and more consistent positioning of CGIAR and its role within the R4D continuum and new metrics on the efficacy of capacity

²¹ Annex 5 of the Synthesis report presents five cluster analyses by five subject matter experts, where recommendations were made for each cluster. Their recommendations were pooled and systematically collated to feed into final recommendations. As such, the recommendations for each of the five clusters can be traced to the recommendations in the main body of the report, although not consistent with their numbering within Annex 5.

development in enabling others to take forward CGIAR's research processes and products for themselves.

- 4** Define CGIAR's comparative advantage in delivery of different elements of the ambitious 2030 Research and Innovation Strategy and its projected scale of funding: review where internal investments and capacities are most needed, and where gaps can be more effectively met through external partnerships.
- 5** Strengthen country and regional coordination structures as a facility for all CGIAR Centers/research initiatives to explore integrative solutions at local, landscape, and relevant subnational, national and regional scales, ensuring coherent and responsive engagement with national stakeholders and agenda. These can leverage assets & scientific knowledge, local relationships & reputation developed by Centers over five decades. The Partnerships review noted the significant cross-CRP efforts made in this regard through GCARD 3 at the start of Phase 2, but the process was not followed through – a lack of coherence is also a burden on NARES partners.
- 6** Operationalize a high quality, common approach to research ethics and science quality and their measurement.
- 7** Maintain effective knowledge management to track processes and findings through successive phases of work and maintain public access to key documents and data. In a number of cases, earlier CGIAR reports were not available on the internet to corroborate elements of this synthesis.
- 8** A policy-level shift both in practices and in evaluation is required from the norm of seeking to attribute development impacts to CGIAR research, towards determining and valuing the essential contribution CGIAR is making with others both through its research, and by mobilizing collective actions among diverse public, private and civil society partners to transform innovation systems for development impact.
- 9** Enhance determination of QoS through bibliometric analyses and facilitate comparison across CRPs and new research initiatives by a) maintaining the same data source over time, b) obtaining citation data annually to enable direct comparisons unaffected by the number of years elapsing, c) retaining data from analyses its raw format, including all metadata, to allow data to be re-analysed in future and be visualized in new ways, and d) developing standard guidance and indicator definitions.

4.2 Recommendations for the Three Strategic Action Areas

4.2.1 Systems Transformation (AA1)

- 10** Rather than tackling climate change, NRM, and agriculture for nutrition/health separately in CGIAR, target research to science-policy engagement as part of a broader, integrating effort on transforming food systems.
- 11** Research design needs to focus more strongly on poverty reduction across all programs; target the rural resource-poor, women and those most disadvantaged. Increase attention to understanding and addressing the equity and impacts of policies, shocks and risks faced by the poor in taking up technologies and research their solutions.
- 12** Address the linkages between environmental sustainability and resilient agri-food systems. Relationships between the dynamics of environment, ecosystems, biodiversity and livelihoods in agro-ecosystems will require significant attention.
- 13** Identify a handful of place-based programs in priority agro-ecologies where the triple challenge of achieving sustainable food production, enhancing human well-being, and conserving ecosystem services can be addressed, and where national commitments bring opportunity for impact at scale through integrated innovation systems.

4.2.2 Resilient Agri-Food Systems

- 14** Reorient research designs to focus more on the vulnerable poor, in particular women and the disadvantaged and those at greatest risk from natural resource depletion, severe climate change impacts, economic deprivation and conflicts.
- 15** Strengthen the capacity to innovate to improve the resilience of agri-food systems, not only to adopt specific technologies. Improve assessment and metrics related to risk and resilience, including innovation capacity, and co-develop social and technical innovations with at-risk populations
- 16** Systems transformation will require coherent development and adoption of technical and social innovations at scale. One CGIAR should give greater emphasis to research on scaling science and implementation science.

4.2.3 Genetic Innovation

- 17** Ensure high priority is given to nutrition/health/resilience/environmental sustainability objectives in research groups focused on genetics.
- 18** Contextualize variety production through greater inclusiveness in defining product profiles, executing programs and delivering outputs, to tailor research to diverse agricultural communities and to the needs of children, youth, women, and other at-risk or marginalized group.
- 19** Put greater priority on seed sector development to facilitate impact at scale, including expanding partnerships with the private and civil society sectors and strengthening key policies and regulations.
- 20** Catalyze partnerships with other research and innovation partners in defined systems to enable crop system diversification and improved access to affordable, healthy diets.
- 21** Accelerate the modernization and technical capacity development of plant breeding programs, across Centers and in National Program partners.
- 22** Integrate research with wider development and investment commitments related to climate change adaptation and mitigation.
- 23** Engage strategically with policies (e.g., ITPGRFA, CGRFA) around the value of germplasm diversity, farmers rights and breeders rights to plant and animal genetic resources and international transfer agreements, to ensure access and availability of diverse and valuable germplasm, improved varieties and strains and crop wild relatives.

4.3 Recommendations for Seven Ways of Working

4.3.1 Embracing a Systems Transformation Approach, Seeking Multiple Benefits across Impact Areas

- 24** Ensure that public, private and civil society stakeholders are involved in foresight and priority setting processes and have a sense of ownership about the research agenda.
- 25** Strengthen the systematic incorporation of equity issues as important dimensions of research design and analysis. Diversify partners and skills to better address root causes of challenges, to include, for example, social scientists, and experts from the private sector, sustainable finance, and humanitarian sectors. Expand socio-economic work including poverty and livelihood assessments, adoption studies, policy and institutional analyses, and in-depth gender and youth studies, with strengthened in-house capacity and/or additional partners.
- 26** Invest in training researchers in systems science and build research from a shared understanding of agri-food systems that integrates objectives related to production, livelihoods, environment/biodiversity, and health/nutrition and takes a holistic approach to AFS and risk management and participatory innovation approaches to engage with farmers and rural communities.

- 27** Strengthen MELIA metrics and user-friendly, streamlined reporting systems based on simple, nested TOCs, developed with, and owned by, partners and stakeholders, with required baselines, actions, capacities and responsibilities coherently planned towards achieving desired outcomes. A special effort should be dedicated to check continuously the validity of assumptions. The role of CGIAR and key stakeholders in the cycles of learning and change from identifying needs to basic research outputs to impact on the ground at the smallholder level needs to be clarified and better captured by the TOCs.
- 28** Tailor corresponding metrics to CGIAR's comparative advantage and realistic expectations of CGIAR's contributions to sustainable development outcomes across the 5 impact areas.
- 29** Incentivize the use of MELIA metrics for progressive cycles of evidence-based learning and adaptive management, working in close collaboration with partners and stakeholders, to optimize delivery and impacts. Increase the use of mixed methods designs in evaluations, with metrics for outcome pathways that go beyond CGIAR and its immediate boundary partners, to bring broader perspectives, a more informed understanding of CRP's effectiveness, and therefore improved decision making. Contribution analysis for such outcomes needs to be determined not only from within CGIAR, but also be recognized and acknowledged from the perspectives of others, whether directly involved or indirectly affected by CGIAR's work. These should be carried out at country or landscape level, considering also contributions of other research and innovation actors involved.
- 30** Improve the coverage of cross-cutting themes (e.g., gender, youth) in MELIA by strengthening MEL disciplinary skills as applied to design and implementation.
- 31** Expand technical assistance on MELIA to research managers, scientists and partners.

4.3.2 Leveraging Ambitious Partnerships for Change

- 32** Take a more systematic approach to partnership development, and to individual and institutional capacity development, at all levels. Develop strategies for partnership and capacity development. Establish explicit timebound targets for progressive transfer of responsibilities and resources to enable local partners to sustainably take on a research/innovation area for themselves through agreed exit strategies.
- 33** Draw more extensively on CGIAR's value as a broker of networked actions in making significantly greater use of research and development partnerships to fill knowledge and skill gaps in the research processes and innovation webs involved, enabling CGIAR to focus on its own strengths and areas of comparative advantage. These should include partnerships with the private sector throughout the food system and with non-CGIAR ARIs, including South-South partnerships, as well as SMEs and CSO agencies for scaling of innovations, value addition and market access. Facilitate partnerships linking non-CGIAR ARIs to local and national partners for collaborative research and capacity development in the new initiatives. Explore opportunities for CGIAR programs to contribute productively to national development agendas, foster synergies and reduce duplication of effort. Established as service providers to CGIAR, the GENEBANK& EIB platforms have the potential to strengthen genetic conservation/use and advanced breeding capabilities in national systems.

4.3.3 Positioning Regions, Countries, and Landscapes as Central Dimensions

- 34** Prioritize responsiveness of the research agendas to local, national and regional strategies and initiatives to facilitate the achievement of outcomes at scale. Initiate or strengthen long-term, transdisciplinary research at dedicated field facilities strategically located in relevant landscapes of developing countries. Co-locate activities from many programs in these geographic areas, to better coordinate outcome-driven research activities, build partnerships and share infrastructure.
- 35** Develop consistent communication policies and practical, ethical guidance to inform CGIAR engagement with local partners at different levels (communities, government, private sector, NGOs, ARIs). Communicating in the right way with local partners is essential; CGIAR should expand its in-house communications and outreach capacities and ensure that country-based

staff are well-trained. Translate systems CRPs and global integrating programs' experiences with developing, funding and managing Platform research initiatives through widening participation of local partners into guidelines that can be used across CGIAR.

4.3.4 Generating Scientific Evidence on Multiple Transformation Pathways

- 36** Strengthen social science capacities to complement the biophysical expertise of CGIAR through increasing in-house resources or external partnerships. Integrate social scientists into action research and develop appropriate incentives to encourage interdisciplinary and systems research.
- 37** Invest in creating a shared vision – including stakeholders and researchers – on what could be achieved in a group of research activities at the region, country, landscape or community level and a TOC on how to achieve change. A successful process will require significant attention to facilitating communications among the different levels of researchers/stakeholders.

4.3.5 Targeting Risk Management and Resilience as Critical Qualities

- 38** Expand work on risk and resilience assessment and managing risk throughout the food system through strengthening CGIAR capacities or engaging external partners. Elevate the priority on improving resilience to climate and pest stresses in the developing/adapting/assessing technologies and innovations for crops and livestock.

4.3.6 Harnessing Innovative Finance

- 39** Collaborate with ARIs and the private sector to explore new financing avenues for expanding research on managing risk to unlock access to finance, inputs and innovation-based enterprise opportunities for women, youth and other marginalized groups, building on index insurance, blended (public-private and public-private-producer) finance models and other new approaches.
- 40** Pursue direct linkage between CGIAR in-country coordinated research for development actions, development loan and grant assistance to countries, as well as direct co-financing where feasible.

4.3.7 Making the Digital Revolution Central to Our Way of Working

- 41** Digital innovation and big data were not systematically assessed in the evaluations used as the base for this synthesis. However, CRPs and platforms have been active in these areas, and the opportunities are enormous. A wholesale review of CGIAR capacities and opportunities around big data and practical field applications for pro-poor sustainable development would involve:
 - Expanding the use of remote sensing and GIS
 - Exploring the ethical applications of artificial intelligence, big data, and citizen science through which the poor can specifically benefit
 - Assessing lessons from the rapidly expanding use of digital tools for breeding, weather and agronomic information, extension, and marketing.

Action on all recommendations would increase the contribution of CGIAR research towards sustainable development impacts. Selected recommendations in Box 3 in the Executive Summary are particularly well aligned towards the five CGIAR Impact Areas.

5 Key Evidence Gaps

The systematic review of evaluative evidence, based on the analytical framework, provides useful information on trends and patterns for some of the themes, in particular QoS outputs, CRP design, and ToC. For other themes, however, evidence was limited, fragmented, or inconsistent to make conclusions about their worth. Nevertheless, certain key gaps identified in phase 1 evaluations and reviews began to be addressed in phase 2 reviews (e.g., outcomes were reviewed through OICRs) as well as through thematic/cross-cutting thematic studies. The relative scale of theme (and subtheme) coverage is shown in Figure 6, which reflects the most frequently assigned codes across the two CRP phases.

Figure 6. Code clouds of the coverage of themes in G1 evaluations, across 2 phases



CRPs-focused reports released before or in 2017 (on the left) and after 2017 (on the right).

5.1 Thematic Gaps

- A. Effectiveness of integrating cross-cutting issues in CRPs programs:** Beyond gender dimensions, which were increasingly taken into account, little evidence was found from which to draw conclusions about the extent to which the needs of youth, the landless, indigenous peoples, and other groups of concern have been specifically addressed by CRP activities. This gap also reflects limited social science capabilities in the CRPs.
- B. Capacity development outcomes:** While the evaluations largely covered immediate capacity development activities and outputs, little evaluative evidence was available by which to determine subsequent outcomes.
- C. Challenges and opportunities with regard to digital innovations:** Evidence for digital innovation in the CRPs has been scant, as this topic was not itself evaluated. A wholesale review of CGIAR capacities and opportunities around big data and practical field applications for sustainable development and equity would bring an important focus for change (see recommendations).
- D. Stakeholders' engagement and CGIAR's role within the AR4D continuum:** Evidence was fragmented and not consistent enough to assess the quality of CRPs' engagement with national public and private partners. Largely, evaluations did not show that intended end users had a strong and equitable voice in prioritizing research or that local public, private, and civil society partners had access to the financial resources, capacities, and enabling measures required to access, transform, adopt, and make use of research products.
- E. CRPs outcomes, impact, and sustainability:** Evaluation reports recognized the need for better assessment and documentation of outcomes and impact. The lack of appropriate MEL for CGIAR's development objectives is a significant handicap in seeking to demonstrate the extent to which CRPs contribute to outcomes or in making recommendations for the future. Evaluations were strong on the Common Results Reporting Indicators (introduced in 2017) (i.e. the number of publications, people trained, and number of innovations produced, but these largely relate to the research output level (CGIAR's "sphere of control").²² The evaluations inconsistently assessed the changes in knowledge, skills, attitudes, and/or relationships that manifest as changes in the behavior of output users to which research outputs and related activities have contributed.

²² CGIAR, *Performance Report 2018: Food Systems Innovations for Impact*, <https://www.cgiar.org/progress-towards-research-outputs-and-outcomes/>.

F. *Lack of evaluation of outcomes:* The lack of follow-through in evaluating outcome measures—such as the use made of capacity development; changes in practices, policies, and institutions; and direct and indirect benefits in CGIAR’s “sphere of influence”—is particularly surprising, given the premise upon which the CRPs were established and funded. OICRs and impact assessments have been helpful in highlighting such achievements but by themselves do not provide a comprehensive assessment of progress. Moreover, the validity of assumptions underlying outcome pathways were scarcely questioned in evaluation reports. When combined with the diversity and inconsistency in outcome measurement across phases and CRPs, these factors make it challenging for synthesis work to draw constructive conclusions on outcome pathways.

5.2 Methodological Gaps

- G. *Context of evaluations:*** Country and regional evaluations were not generally undertaken by IEA/CAS and CRPs themselves. Moreover, there was a methodological gap in that local stakeholders’ voices were largely absent from 2020 reviews. The methodological design was fully desk-based, mainly because of the travel restrictions caused by the COVID pandemic, but a wider framing could have still solicited responses from other partners from intended outcome pathways. Phase 1 evaluations used case-based designs including country field visits, yet local socioeconomic contexts, opportunities, and risks were still not brought out strongly at the higher level and were generally lost from the summaries and recommendations sections. There is a real risk of underestimating the influence of local socioeconomic and socio-political factors on achieving intended outcomes and contextualizing CGIAR research accordingly. This gap needs to be addressed in future evaluations.
- H. *Metrics on CGIAR mission:*** The dearth of metrics tailored to CGIAR’s mission reduced the value and consistency of evidence gathered and analyzed in the evaluations. Evaluators had limited time, and their analyses could have benefited from reliable monitoring and evaluation, knowledge management systems, and databases. Metrics should be urgently developed and applied to determining CGIAR’s contributions (IDOs and SLOs) to sustainable development outcomes across the full range of intended impacts.
- I. *Cross-cutting themes:*** There was a lack of depth in assessing cross-cutting themes—e.g., gender, youth, climate change, NRM, and capacity development outcomes. Moreover, few studies adopted an inclusive approach, such as ensuring that the evaluation process was informed by gendered inequality issues, undertaking subgroup analysis, or using appropriate methods to understand differential effects for youth and vulnerable groups.
- J. *Phase 2 assessments:*** Lack of assessment of specific topics in the second phase, such as the GENDER and GENEBANK platforms, constrained the comprehensive analysis of progress along and between phases, as well as along CGIAR strategies and frameworks.
- K. *Contribution analysis:*** Methods remain skewed towards attributing results and outcomes to CGIAR’s work. Given that CGIAR is only one actor within the complex reality of innovation webs, there is a strong need for more contribution analysis, using mixed-method approaches to determine, from the data and perspectives of others, the scale and value of the contribution made by CGIAR’s work toward achieving sustainable development outcomes. This work will also help identify CGIAR’s specific comparative advantages and areas where others would be better placed to deliver required actions.



Advisory
Services

CGIAR Advisory Services (CAS)

Via di San Domenico 1, 00153 Rome, Italy

Email: cas@cgiar.org

URL: <https://cas.cgiar.org/>