Research Article Open Access

# Building a Future-Oriented Breeding CRO Service Platform: Development Pathways for Standardization, Compliance, and Intelligence

Xuanjun Fang, Qixue Liang
Hainan Provincial Key Laboratory of Crop Molecular Breeding, Hainan Institute of Tropical Agricultural Resources (HITAR), Sanya, 572025, Hainan, China
Corresponding email: xuanjunfang@hitar.org
Molecular Plant Breeding, 2026, Vol.17, No.1 doi: 10.5376/mpb.2026.17.0001
Received: 04 Jan., 2026
Accepted: 02 Feb., 2026
Published: 15 Feb., 2026

**Preferred citation for this article**:
Fang X.J., and Liang Q.X., 2026, Building a future-oriented breeding CRO service platform: development pathways for standardization, compliance, and intelligence, Molecular Plant Breeding, 17(1): 1-17 (doi: 10.5376/mpb.2026.17.0001)

**Abstract** With the integration of molecular breeding techniques and digital platforms, the global breeding ecosystem is undergoing a fundamental shift-from fragmented outsourcing services to platform-based, intelligent collaboration. As a critical interface connecting research institutions, enterprises, and regulators, breeding CROs (Contract Research Organizations) are evolving from experimental executors into integrated service hubs characterized by standardization, regulatory compliance, and AI-enabled intelligence. This paper proposes a triadic capability framework encompassing standardized workflows, full-spectrum compliance governance, and intelligent system integration. It diagnoses structural challenges such as the lack of service standards, regulatory mismatch, data fragmentation, and shallow application of AI tools. Furthermore, the study outlines actionable strategies for platform development, including modular SOP libraries, embedded compliance pipelines, data-driven AI middle platforms, and collaborative visualization dashboards. The paper concludes by envisioning an AI-powered transformation of breeding services and recommends advancing institutional pilots, platform certification standards, and AI governance mechanisms to establish CROs as credible, intelligent, and open infrastructure in global breeding innovation.
**Keywords** Breeding CRO; Standardized services; Intelligent platform; Regulatory compliance; Artificial intelligence; Data interoperability; Digital breeding infrastructure

With the deep integration of molecular breeding and digital technologies, the global bio-breeding system is undergoing a profound transformation from tool integration to platform-based collaboration. Its core characteristics can be summarized as the coordinated evolution of “molecularization, intelligence, and systematization”. On the one hand, the continuous decline in the cost of genome sequencing and multi-omics technologies has promoted molecular design breeding as a key paradigm for new variety development. Molecular marker technologies represented by SNPs and genomic selection models have significantly improved the efficiency of trait identification and genetic improvement (Xu et al., 2017). On the other hand, breakthroughs in gene-editing technologies such as CRISPR/Cas have provided efficient tools for the targeted modification of functional genes, substantially accelerating the creation of elite traits (Razzaq et al., 2021). Meanwhile, the widespread application of remote sensing, sensors, and high-throughput phenotyping platforms has enabled large-scale and real-time acquisition of phenotypic data. Driven by artificial intelligence and big data, breeding is now entering an intelligent stage (Zhu et al., 2024).

To address the growing complexity of research and development (R&D) processes, as well as increasing pressures in quality control and regulatory compliance, the Contract Research Organization (CRO) model has begun to see accelerated application in the field of bio-breeding. Originating in the pharmaceutical industry, the CRO model represents a specialized service structure that enhances R&D efficiency and reduces compliance risks for innovation entities through standardized procedures, professional teams, and auditable systems. Existing studies have shown that breeding CROs can effectively resolve coordination inefficiencies in traditional breeding systems caused by fragmented resources, non-standardized processes, and regulatory pressures, and have become an important supporting force in modern breeding systems (Fang and Liang, 2026).

Building upon the prior systematic summary of the evolutionary pathways and platform models of breeding CROs (Fang and Liang, 2026), this study further focuses on constructing a future-oriented service platform architecture.

While breeding CRO platforms have developed rapidly, they have also exposed structural challenges, including the absence of unified service standards, limited regulatory adaptability, and unclear pathways for digital transformation. These issues call for breakthroughs in institutional design and systemic capabilities.

This study focuses on three key directions for the future development of breeding CROs—standardization, compliance, and intelligence—with the aim of systematically proposing a theoretical framework and practical pathways for platform-based CRO services:
Standardization: Establish unified Standard Operating Procedure (SOP) and Good Laboratory Practice (GLP) systems across species and application scenarios to enhance service consistency and data reproducibility (Liang and Zhou, 2012);
Compliance: Strengthen regulatory alignment and risk control capabilities in areas such as transgenic technologies, gene editing, biosafety, and data governance;
Intelligence: Integrate AI models with high-throughput experimental systems to build data-driven intelligent decision-making platforms, thereby improving the efficiency and responsiveness of breeding services.

Through these pathways, this study aims to provide a replicable structural framework to support the transformation of breeding CRO platforms from "technology-execution-oriented" entities to "system-empowerment-oriented" platforms, and to offer theoretical support and practical demonstration for quality governance, resource allocation, and technological transformation in the modern seed industry.

# 1 Industry Status and Challenge Diagnosis

## 1.1 Lack of standardization undermines service mutual recognition

The breeding CRO industry is currently at a developmental stage characterized by "multiple emerging entities operating independently". Most institutions lack unified standards in experimental procedures, data collection, field trial management, and quality control systems. Significant differences exist in Good Laboratory Practice (GLP) and Standard Operating Procedure (SOP) systems, making it difficult to compare, reuse, and reproduce service outcomes across institutions (Liang and Zhou, 2012; Van Etten et al., 2023). This fragmentation not only constrains the professional development of breeding CROs but also limits their integration into domestic or international regulatory and certification frameworks, thereby affecting high-end clients' confidence in data credibility (Lassoued et al., 2018; Menz et al., 2020).

In contrast, institutions such as the United States Department of Agriculture (USDA) and the European Food Safety Authority (EFSA) have implemented clear quality standards and regulatory interfaces in outsourced agricultural R&D services. CRO services in Europe and the United States generally comply with systems such as ISO 17025 and OECD GLP, providing a foundation for cross-border mutual recognition of data. In comparison, a dedicated industry-wide standard system for breeding services has yet to be established domestically, and this lack of standardization has become a critical bottleneck restricting the upgrading of breeding CROs.

## 1.2 Weak compliance systems generate high risks

At present, most breeding CROs have not established a systematic compliance management mechanism covering the entire process of "pre-trial–in-trial–post-trial". Compliance risks are particularly significant in areas such as genetically modified organisms (GMOs), gene-edited materials, biosafety, and material transfer. Standardized procedures for interfacing with regulatory authorities (e.g., the Ministry of Agriculture and Rural Affairs, USDA, and EFSA) are often lacking, as are institutional mechanisms governing non-disclosure agreements (NDAs), material transfer agreements (MTAs), and intellectual property (IP) allocation (Purnhagen and Wesseler, 2020; Qaim, 2020).

Compliance challenges become even more pronounced in international collaborative projects. Regulatory approaches to new breeding technologies vary considerably worldwide: the United States generally adopts a product-based regulatory framework, whereas the European Union emphasizes a process-based regulatory approach and maintains particularly stringent oversight of GMOs (Davison and Ammann, 2017). These inconsistencies in regulatory pathways hinder cross-regional mutual recognition of trial data and increase both the regulatory interpretation costs and operational complexity of CRO services (Menz et al., 2020; Qaim, 2020). Therefore, establishing a compliance system aligned with both domestic and international regulatory frameworks is a critical prerequisite for breeding CROs to sustain participation in international collaboration.

### 1.3 Data silos limit the release of intelligent potential

With the rapid accumulation of multi-dimensional data—including phenomics, genomics, and enviromics—the "data organization capability" of breeding CRO platforms is gradually replacing traditional "experimental execution capability" and becoming a core competitiveness. In practice, however, molecular testing data, phenotypic data, trial management records, and compliance documentation are often dispersed across different systems. These datasets vary in standards and formats, making cross-platform sharing and reuse difficult (Fernandez et al., 2020; Mahmood et al., 2022).

In the absence of unified data interfaces and platform architectures, most CROs are unable to support the continuous iteration of machine learning models or conduct large-scale data training across years and crop species. This limitation directly constrains the practical implementation of intelligent functions such as AI-assisted breeding design, trait prediction, and trial optimization (Yan and Wang, 2022; Van Etten et al., 2023).

Leading international institutions have begun to establish integrated platforms that unify data acquisition, quality control, compliance documentation, and client interfaces within a single system architecture. For example, CGIAR in the United States and the EJP Soil program in the European Union are promoting interoperability among agricultural data platforms to enhance data openness and reuse value. These experiences indicate that building an open, standardized, and intelligent data infrastructure is a critical direction for the digital transformation of breeding CROs in the future.

### 1.4 Ambiguous terminology and regulatory boundaries affect industry recognition

At present, the term "breeding CRO" lacks a consistent definition both domestically and internationally, making it difficult to clearly distinguish it from general technical service outsourcing, public breeding platforms, and trial contracting institutions. In both academia and industry, no consensus has yet been reached on key issues such as the division of responsibilities between "contract research" and "collaborative trials", data ownership, and intellectual property management (Lassoued et al., 2018).

At the international level, agencies such as the United States Environmental Protection Agency (EPA) and the USDA have established explicit qualification requirements, data usage standards, and reporting format specifications for outsourced services in areas such as pesticides and genetically modified crops. However, in the field of bio-breeding, where emerging technologies and regulatory frameworks are still evolving, the role of CROs has not yet been systematically incorporated into relevant regulatory structures. As a result, breeding CROs are often overlooked in terms of policy support, qualification recognition, and public funding allocation, which in turn weakens client awareness and trust in their role (Qaim, 2020; Van Etten et al., 2023).

Therefore, establishing a clear terminology system and regulatory interface framework is a prerequisite for the transition "from fragmented service provision to a platform-based industry structure".

## 2 Platform Development Strategies: Reshaping the Capability System of Future-Oriented Breeding CROs

### 2.1 Core platform architecture: a problem-oriented "trinity" capability framework

Problems and Challenges: existing breeding CRO platforms commonly suffer from fragmented capabilities and disconnected processes. Standardization, compliance, and intelligence are often developed independently rather than through systematic integration, making it difficult for platforms to support multi-project operations and cross-regional collaboration.

Development Pathway: The future-oriented breeding CRO platform should build a "trinity" capability framework of "standardized service process—full chain compliance management—intelligent system integration" on the top-level architecture (Figure 1). These three capabilities should evolve synergistically within a unified platform rather than being linearly aggregated (Ezzelle et al., 2008; Smulders et al., 2021; Xu et al., 2022).

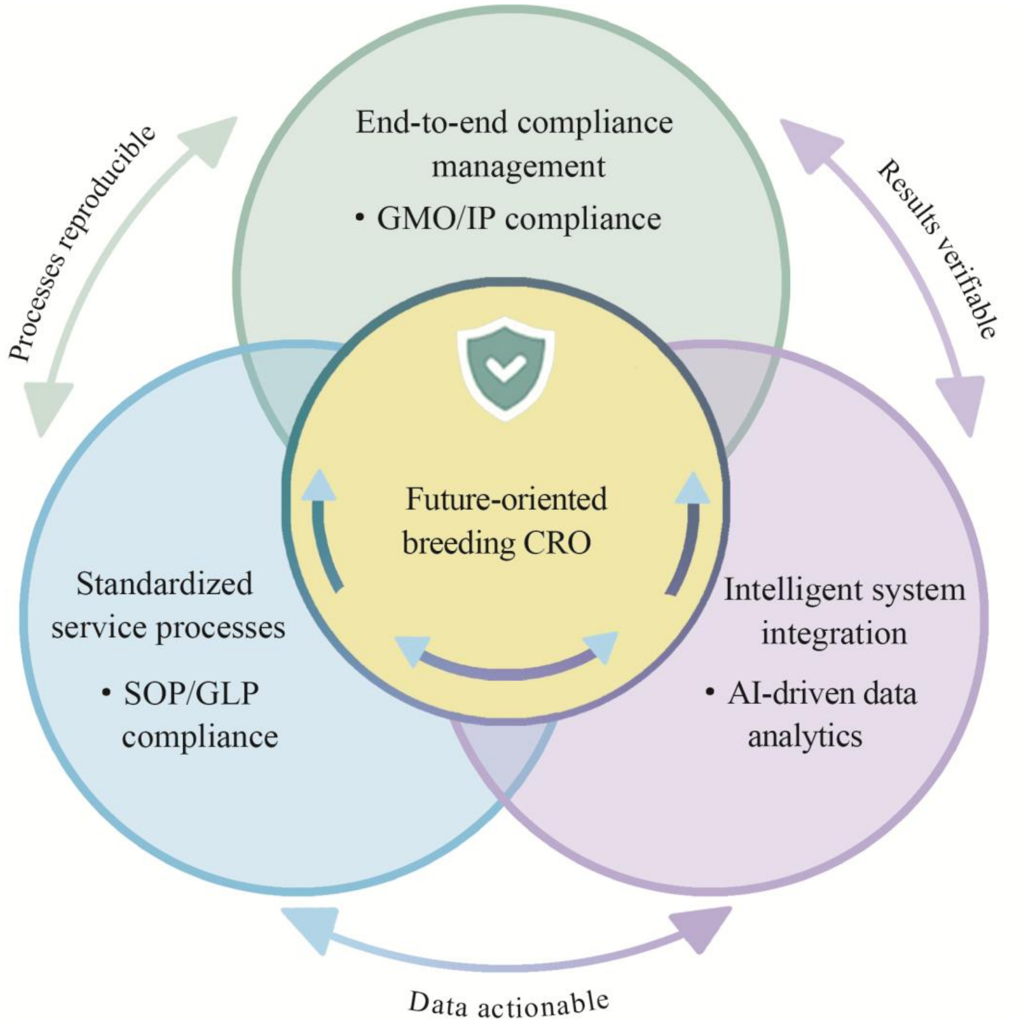


Figure 1 A triadic structural model of the breeding CRO capability system

Figure caption: An integrated triad of capabilities for future-oriented breeding CRO platforms. The diagram illustrates three essential capability modules for breeding CROs: standardized service processes (e.g., SOP/GLP compliance), end-to-end regulatory compliance (covering GMO/IP standards), and intelligent system integration (AI-driven data analytics). These elements interact to form a closed loop of standardization, compliance, and intelligence, enabling data-driven, auditable, and scalable operations across diverse breeding projects and scenarios

This framework emphasizes three key principles:
(1) Standardization as the foundation, addressing the issue of non-reusable processes;
(2) Compliance as the boundary, addressing the issue of non-auditable results;
(3) Intelligence as the amplifier, addressing the issue of data that cannot be transformed into actionable decisions.

Together, these three dimensions constitute the core capability combination that distinguishes platform-based breeding CROs from traditional technical outsourcing institutions.

### 2.2 Standardization capacity building: from fragmented procedures to replicable service modules

Problems and Challenges: although many breeding CROs have established SOP or GLP systems, these are often “project-customized” in nature and difficult to reuse across different crops, teams, or regions. As a result, standardization has not been effectively translated into scalable operational capacity.

Development Pathway: standardization efforts should shift from a "document-oriented" approach to a "module-oriented" approach by constructing a combinable and iterative SOP module library centered on key nodes of the breeding process (Figure 2).

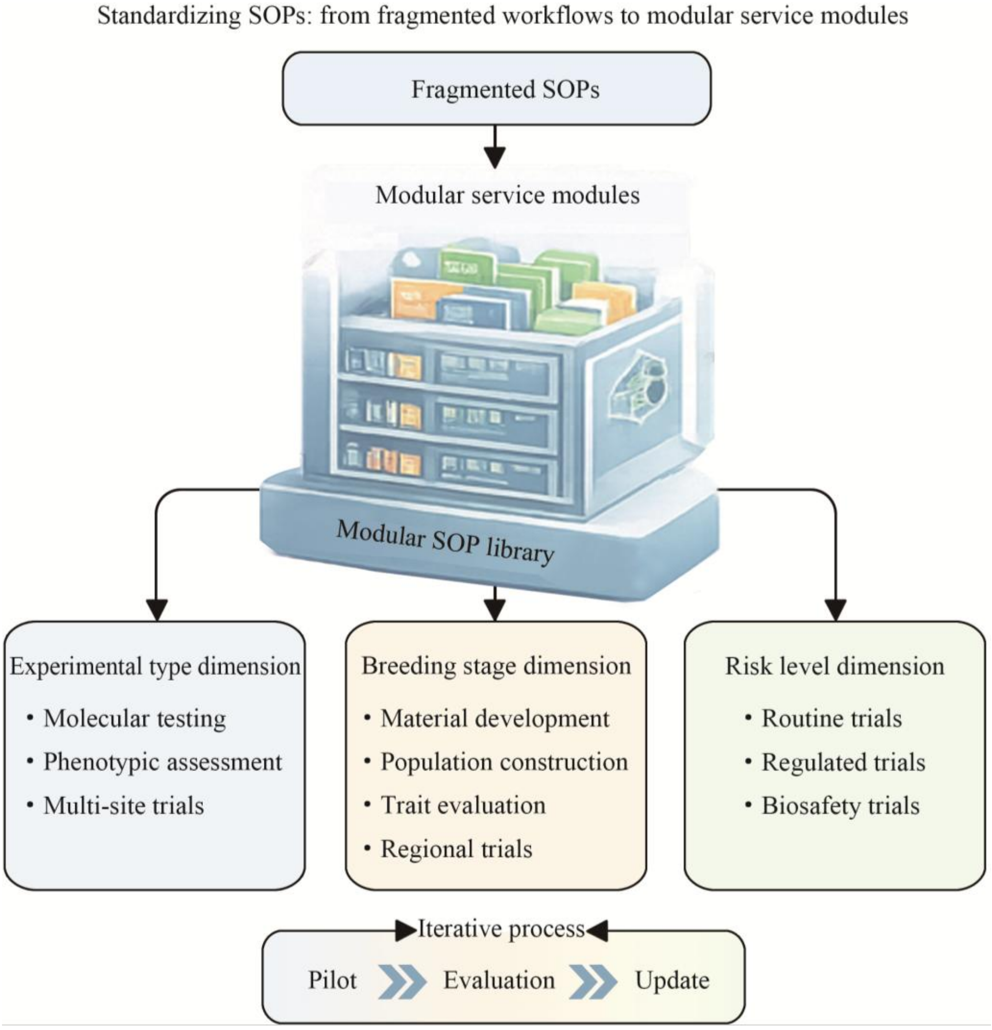


Figure 2 Modular pathway for standardized capability development in breeding CROs

Figure caption: This figure illustrates the transformation of breeding CRO standardization from fragmented, project-specific SOPs to a modular and reusable service framework. By consolidating individual SOPs into a modular SOP library and abstracting them across breeding stages (material development, population construction, trait evaluation, and regional trials), experimental types (molecular testing, phenotypic assessment, and multi-site trials), and risk levels (routine, regulated, and biosafety-related trials), standardized workflows are converted into scalable service modules; The iterative cycle of *pilot–evaluation–update* highlights the role of SOPs as a "living document system," enabling continuous improvement, quality assurance, and cross-project reproducibility within platform-based breeding CROs

At the platform level, SOPs should be abstracted and structured along the following dimensions rather than repeatedly enumerating specific operational details:

Breeding stage dimension: material development, population construction, trait evaluation, and regional trials;

Experimental type dimension: molecular testing, phenotypic assessment, and multi-site trials;

Risk level dimension: routine trials, regulated trials, and biosafety trials.

Through an iterative mechanism of "pilot – evaluation – update", SOPs can evolve from one-time regulatory documents into a dynamic "living document system". When integrated with the quality management system, this approach enables closed-loop management of execution, deviation monitoring, and continuous improvement (Kendall et al., 2016; Gumba et al., 2018a).

### 2.3 Compliance capacity building: from reactive response to embedded governance

Problems and challenges: at present, compliance systems in breeding CROs are largely concentrated on GMO-related projects and often rely on manual, experience-based judgment. Such approaches are insufficient to address systemic risks arising from cross-jurisdictional collaboration, cross-border data transfer, and complex intellectual property (IP) arrangements.

Development pathway: compliance capacity should be upgraded from an "externally imposed requirement" to an "embedded platform mechanism". Through the coordinated integration of institutional design, procedural control, and digital technology, compliance outputs can become standardized and replicable.

At the institutional level, it is recommended that the platform establish a dual-layer structure consisting of a general compliance workflow and scenario-specific pathways:
General framework: source verification—risk classification—trial approval—environmental monitoring—result archiving;
Scenario adaptation: load regulation based on target markets (China, the United States, and the European Union) (Turnbull et al., 2021; Mu et al., 2025).

At the operational level, NDAs, MTAs, and IP contracts should be linked to specific trial process nodes. Through digital systems, access rights controls, audit trails, and document generation can be automatically triggered, thereby reducing uncertainty associated with manual intervention (Figure 3) (Tekic et al., 2023).

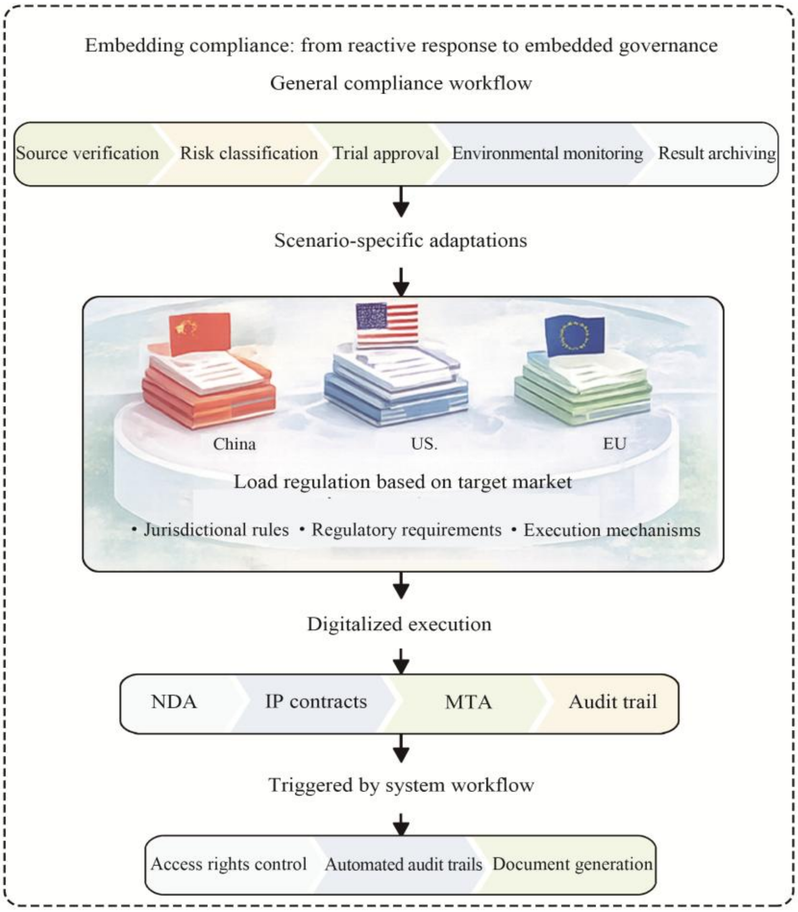


Figure 3 Embedding compliance: from reactive response to embedded governance

Figure caption: This diagram illustrates the transformation of compliance management in breeding CRO platforms from reactive response models to embedded governance systems. It highlights a dual-layered structure-generic compliance workflows and scenario-specific adaptations—supported by digital execution mechanisms including automated audit trails, permission control, and document generation

### 2.4 Intelligence capacity building: from data accumulation to intelligent decision support

Problems and challenges: although breeding CROs have accumulated large volumes of molecular, phenotypic, and environmental data, these datasets are often structurally fragmented and difficult to reuse across projects. As a result, intelligence development frequently remains at the level of "tool adoption" rather than systemic transformation.

Development pathway: intelligence capacity building should center on a platform-level data middle platform rather than isolated AI applications.

The platform should prioritize three foundational tasks:
(1) Establish a unified data model to enable structured integration of genotype-phenotype-environment data;
(2) Introduce AI analytical interfaces for trait prediction, combination optimization, and trial design recommendation;
(3) Deploy visual dashboards to transform analytical outputs into interpretable and decision-oriented information (Han et al., 2020; Copland et al., 2024).

On this basis, breeding CROs can evolve from "experimental executors" into "intelligent decision partner", significantly enhancing their strategic position within collaborative systems (Sumathi, 2025). As illustrated in Figure 4, intelligence capacity building can achieve the transition from data accumulation to intelligent decision-making through the coordinated development of a data middle platform, integration of AI analytical interfaces, and linkage with visualization dashboards.

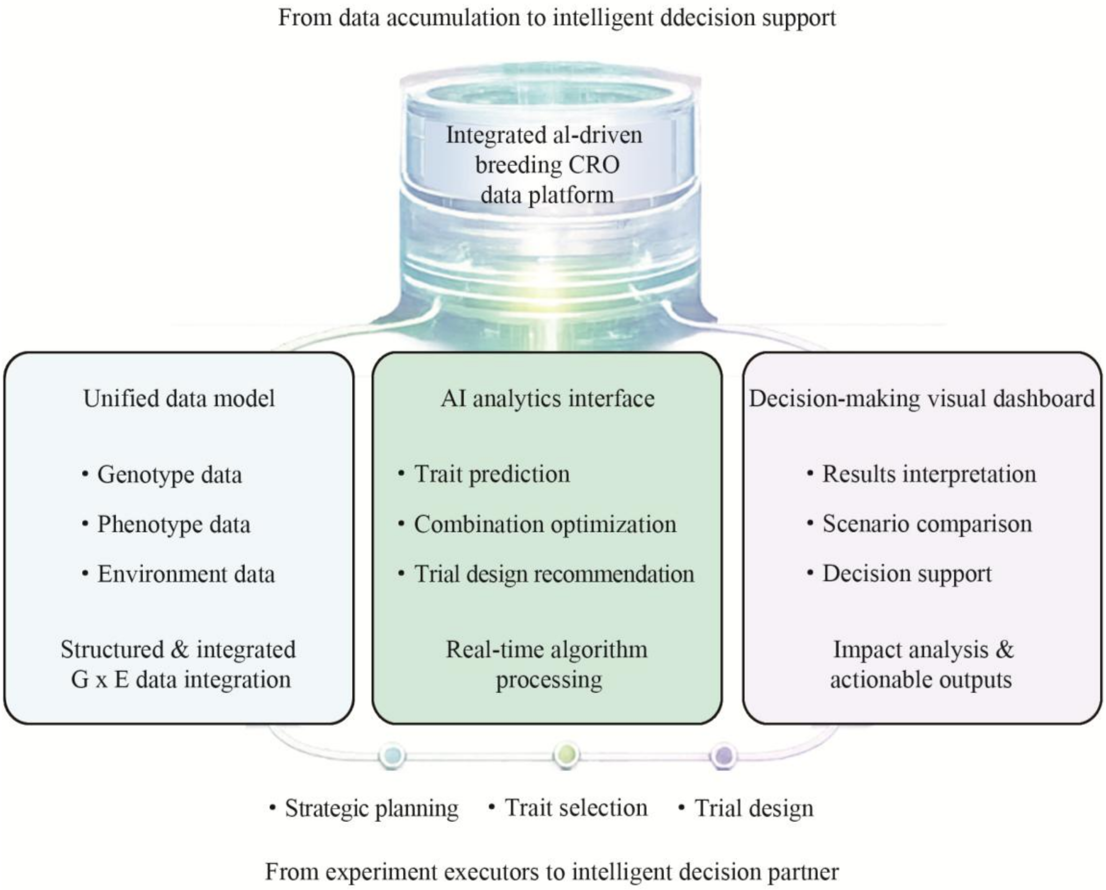


Figure 4 Pathway for developing AI capabilities: from data accumulation to intelligent decision support

Figure Caption: This figure illustrates the pathway for developing intelligent capabilities in breeding CRO platforms. It emphasizes the transformation from fragmented data integration, standardization, and modeling to the deployment of AI interfaces for trait prediction, optimization, and trial design. Finally, the use of visual dashboards enables actionable decision support. This framework empowers breeding CROs to shift from data collectors to intelligent decision enablers

### 2.5 Maturity model for platform

To prevent platform development from remaining at the level of general principles, this study proposes a maturity model for breeding CRO platforms (Table 1) to guide phased construction and systematic evaluation.

Table 1 Maturity model for breeding CRO platforms

| Maturity Level | Standardization Capability | Compliance Capability | Intelligent Capability | Platform Feature |
|---|---|---|---|---|
| Level 1 Initial Stage | Fragmented SOPs | Manual Compliance | Isolated Data Analysis | Experience-Based Operations |
| Level 2 Defined Stage | Basic SOPs | Streamlined Compliance | Centralized Data Storage | Single Project Support |
| Level 3 Integrated Stage | SOPs+GLP Compliance | Embedded Compliance Workflow | Data-Driven Decision Hub | Concurrent Project Management |
| Level 4 Intelligent Stage | Cross-Institutional SOPs | Cross-Jurisdiction Compliance | AI-Enabled Decision Support | International Service Platform |

The maturity model for breeding CRO platforms illustrates the evolutionary trajectory of platform capabilities across three dimensions: standardization, compliance, and intelligence. The four maturity levels—from “Initial” to “Intelligent”—reflect the systematic enhancement of service processes, data management, and platform governance capacity. This model facilitates the assessment of the relative positioning of different breeding CROs in terms of service systematization, digitalization, and internationalization. It provides a clear roadmap addressing “where the platform starts, where it is heading, and how progress can be evaluated”.

### 2.6 Summary: from capability aggregation to system evolution

By restructuring platform architecture through a problem-oriented approach and integrating standardization, compliance, and intelligence into a unified capability system, breeding CROs can transition from “project-based services” to “platform-based infrastructure”. This transformation not only responds to current industry fragmentation and compliance pressures but also lays the foundation for breeding CROs to assume higher-level roles within the global bio-breeding innovation ecosystem.

## 3 Integration of Intelligence and Digital Platforms: Reshaping Services from Algorithms to Systems

### 3.1 Functional evolution of AI in the breeding service chain

As bio-breeding enters an era driven by multi-omics technologies and data explosion, the role of artificial intelligence (AI) in breeding services is evolving from a “single-point tool” to a “decision engine”. AI now spans the entire service chain, from sample collection and phenotypic analysis to complex trait modeling and breeding pathway optimization (Xu et al., 2022; Zhu et al., 2024). To ensure the genuine integration of AI into service platforms, technological iteration alone is insufficient. AI systems must be deeply coupled with existing information systems—such as Laboratory Information Management Systems (LIMS), Electronic Laboratory Notebooks (ELN), and phenotypic recognition systems—to establish a complete data closed-loop and feedback mechanism (Figure 5).

(1) From “data acquisition” to “structured modeling”: integration pathways of intelligent phenotyping systems

In breeding services, phenotypic data collection has long relied on manual operations, resulting in high subjectivity and limited reproducibility. In recent years, phenotypic recognition platforms based on deep learning and sensor networks (e.g., OpenPheno and PhenoBox) have enabled automated measurement of multiple crops under diverse environmental conditions (Ampatzidis and Partel, 2019; Hu et al., 2025). By integrating unmanned aerial vehicle (UAV) multispectral imaging, ground-based rail systems, and environmental sensor networks, and connecting these devices to internal platform systems such as LIMS and ELN, a standardized data pipeline of “acquisition – storage – modeling – feedback” can be established.

For example, the TraitMill platform (BASF) directly transmits sensor-collected data—such as canopy structure, leaf area, and disease indices—to a centralized analytical middle platform, enabling parallel modeling and differential analysis of multiple traits. This integration model, built on standardized interfaces, ensures that phenotypic data not only support real-time analysis but also serve as stable input sources for AI model training.

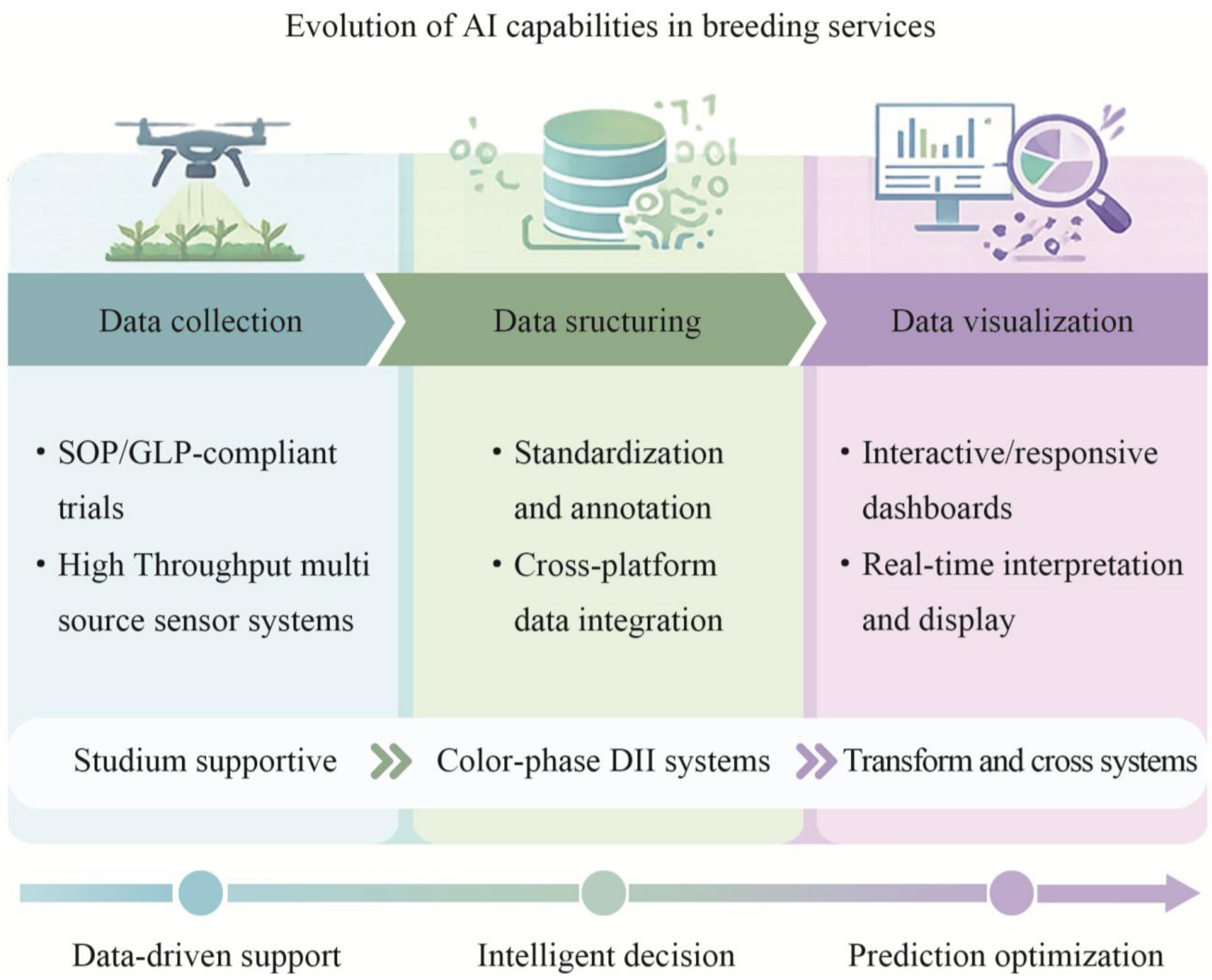


Figure 5 Evolution of AI capabilities in breeding services
Figure Caption (APA format): This figure illustrates the evolution of AI capabilities in breeding services, highlighting a progressive pathway from data collection (e.g., SOP/GLP-compliant trials and multi-source sensing systems), through data structuring (standardization, annotation, and cross-platform integration), to data visualization (interactive dashboards and real-time displays). The bottom sequence demonstrates how AI capabilities transition from data-driven support to intelligent decision-making and predictive optimization—serving as a foundational framework for building future-oriented breeding CRO platforms

(2) From “trait prediction” to “multi-objective optimization”: building a decision-support engine
Complex traits—such as yield and stress resistance—are influenced by polygenic effects and genotype-by-environment interactions, making them difficult to address using traditional analytical methods. AI models, particularly ensemble learning approaches, graph neural networks, and explainable AI techniques, have demonstrated the capacity to perform joint modeling and optimization ranking of high-dimensional traits (Cheng and Wang, 2024; Zhou et al., 2024).

Within Corteva’s Enterprise Breeding System (EBS), AI models are used to evaluate the expected performance of tens of thousands of breeding materials across diverse environments in real time, providing optimal combination recommendations for each selection cycle. Such systems are often integrated with Geographic Information System (GIS) platforms, incorporating field-level indicators such as soil properties, moisture conditions, and pest and disease risk factors. This integration enables a truly three-dimensional “environment–genotype–phenotype” fusion model.

Moreover, in maize breeding programs targeting smallholder farming regions, federated learning models have enabled collaborative data modeling across multiple low-resource countries. This approach has significantly improved selection efficiency for traits related to marginal environmental adaptability, demonstrating the adaptability of AI under conditions of globally uneven data distribution.

(3) From “breeding pathway simulation” to “strategic design”: enabling forward-looking decision-making
Traditional breeding programs often lack dynamic mechanisms for evaluating the long-term effects of strategic decisions. Contemporary AI systems now enable the construction of “in silico breeding” environments, where genomic estimated breeding values (GEBV), multi-generational simulations, and constraints for maintaining genetic diversity are integrated to generate intelligent recommendations for parental selection, crossing schemes, and generation advancement strategies (Farooq et al., 2024; Zhu et al., 2024).

Such systems can be integrated with the front-end experimental design modules of breeding CRO platforms and linked to trial databases and phenotypic feedback systems, forming a closed-loop structure. This enables full-process support spanning “simulation–execution–adjustment” and establishes a continuous optimization cycle grounded in real-world feedback.

### 3.2 Deep integration of platform architecture and AI systems

The development of platform intelligence is often constrained by the problem of “algorithm silos”. To address this limitation, AI should be embedded into existing management systems, operational workflows, and user interaction interfaces, thereby forming a fully integrated service chain.

(1) Modular design: supporting multi-crop and multi-task heterogeneous configurations
To address differences in practical application needs, the platform should adopt a microservices architecture to deploy distinct AI modules—such as phenotypic recognition, trait prediction, and simulation-based recommendation systems—and enable flexible invocation across different crops and projects through containerized deployment (Varshney et al., 2016; Zhao et al., 2022).

For example, the “Seed Breeding Cloud Platform” decomposes molecular marker analysis, field image processing, and AI-based recommendation functions into reusable modules. Within the user configuration interface, multi-task workflows can be rapidly constructed using a flowchart-style “building block” assembly approach (Zhu et al., 2024). This model effectively lowers the entry threshold for platform services and improves cross-project replication efficiency.

(2) Standardized interfaces and system interoperability: connecting nodes across the ecosystem
LIMS, ELN, sensor systems, and phenotypic recognition platforms are often developed by different vendors and operate under heterogeneous data standards. Breeding CRO platforms should therefore establish an API framework aligned with international standards such as BrAPI and MIAPPE, unifying invocation logic and reducing system integration costs (Sempéré et al., 2019).

For example, OpenPheno achieves seamless connectivity with external phenotyping platforms and data middle platforms through BrAPI-based interfaces. Similarly, the Enterprise Breeding System (EBS) integrates with enterprise platforms such as SAP and ArcGIS, enabling project management, environmental data, and financial control to operate within a unified system environment.

(3) User-centered collaborative visualization platforms
Algorithmic capability generates value only when it is effectively utilized by end users. Therefore, breeding CRO platforms should provide visualization dashboards combined with collaborative operation interfaces, enabling real-time trial progress tracking, environmental response analysis, and graphical presentation of model prediction outputs (Zhao et al., 2022).

Platforms such as OpenPheno and PhenoApp have already incorporated functional components including heat maps, timeline views, and GIS layer overlays, supporting result sharing and online discussion among project members. This design is particularly well suited to multi-location, multi-role breeding project teams. Under well-controlled permission management and data security frameworks, it facilitates efficient cross-regional collaboration while maintaining data integrity and confidentiality.

### 3.3 Building AI-driven "systemic platform service capability"

The development of intelligent platforms should not be viewed merely as a technological upgrade, but as a systemic transformation of the service paradigm of breeding CROs. A mature platform must simultaneously possess the following three categories of capabilities:

| Capability Type | Core Functions | Example Platforms / Technologies |
|---|---|---|
| Data Collection & Management | Automated sensor-based data acquisition, LIMS management, metadata standardization | TraitMill, OpenPheno |
| Decision Analysis Support | Multi-trait prediction, genetic gain simulation, G×E modeling and strategy recommendations | EBS, CropGPT |
| User Collaboration Interface | Visual dashboards, real-time logs, role-based access control, cross-institutional project collaboration | Jinzhong Cloud, PhenoApp |

In the future, breeding CRO platforms should follow a pathway of data interconnectivity, model-driven operation, and service collaboration to build an intelligent hub characterized by continuous learning capability, platform openness, and international adaptability. Through this transformation, breeding CROs can move beyond the traditional model of "experimental outsourcing" and evolve toward the co-construction of an intelligent breeding ecosystem.

## 4 Conclusion: Future Pathways for Platform-Based Breeding CROs

### 4.1 Role transformation of CRO platforms: from service outsourcing to breeding infrastructure

With the explosion of multi-omics data, increasing experimental complexity, and the rapid adoption of digital tools, breeding CROs are evolving from traditional "outsourcing service providers" into core nodes within breeding systems. In the future, breeding CROs characterized by platformization, intelligence, and high compliance standards will function as "digital infrastructure" and "collaborative innovation hubs" within the global seed industry innovation system. Their role will extend beyond experimental execution and data analysis to connecting research institutions, enterprises, regulatory authorities, and international partners, thereby supporting complex breeding projects involving multiple environments and stakeholders worldwide (Xu et al., 2022; Zhu et al., 2024). Under this positioning, CRO platforms must establish end-to-end capability closed loops encompassing data integration, intelligent decision-making, trial execution, and compliance support. While accelerating genetic gain, such platforms will also provide a robust technological foundation for global food security and sustainable agriculture.

### 4.2 Institutional support and industry standards: building a trustworthy operational foundation

Although breeding CROs continue to evolve technologically and organizationally, their long-term sustainable development remains highly dependent on institutional frameworks and industry standards. The sector currently faces widespread challenges, including fragmented processes, inconsistent quality standards, and incompatible data interfaces, which significantly constrain platform interoperability and service scalability (Brookes and Smyth, 2024; Panwar et al., 2025). Addressing these issues requires action along two dimensions. First, at the industry level, unified SOP repositories, standardized data structure specifications, and harmonized quality control indicator systems should be established to promote service standardization and process transparency. Second, at the regulatory level, a "breeding CRO regulatory sandbox" mechanism could be introduced to clarify data compliance boundaries, define rules for AI tool usage, and establish platform certification systems, thereby providing a controlled environment and institutional safeguards for innovation (Alexander et al., 2023; Goktas and Grzybowski, 2025). Only through coordinated policy guidance and industry collaboration can a trustworthy ecosystem characterized by high quality, auditability, and mutual recognition be achieved.

### 4.3 AI-driven capability leap: toward an era of intelligent collaborative breeding

The core engine of future breeding CRO development will be intelligent tool systems represented by artificial intelligence. From phenotypic recognition and trait prediction to genetic optimization and virtual breeding pathway simulation, AI is becoming deeply embedded across all stages of breeding services (Zhou et al., 2024). However, the real value of AI can only be realized when it is deeply integrated with foundational infrastructures such as LIMS, ELN, and sensor systems, and embedded within user decision-making processes. Accordingly, CRO platforms must move beyond the mere application of AI tools toward AI-driven platform construction, establishing an intelligent system characterized by bidirectional iteration between data and models, and deep interaction between algorithms and experimental workflows. On this basis, a full lifecycle AI governance framework should be developed, encompassing algorithm interpretability, model fairness, accountability in decision-making, and ethical boundaries, to ensure the sustainability and trustworthiness of technological applications (Shahriar et al., 2023; Al-Kfairy et al., 2024).

Future research may further examine the strategic role of CRO platforms within multi-stakeholder collaborative breeding mechanisms, particularly their functions in data sharing, cross-border regulatory coordination, and the construction of joint innovation networks.

## Authors' contributions

Xuanjun Fang and Qixue Liang were the executors of this study, responsible for literature review and data analysis, as well as the drafting and revision of the manuscript. Both authors read and approved the final manuscript.

## Acknowledgments

The authors sincerely thank the two anonymous peer reviewers for their careful and thorough review of this manuscript and for their valuable comments and suggestions. The authors also extend their gratitude to Ms. Chunyan Tan from the Hainan Institutes of Biotechnology for providing image post-processing (PS) services for the manuscript.

### References

Alexander C., Yarborough M., and Smith A., 2023, Who is responsible for "responsible AI"?: navigating challenges to build trust in AI agriculture and food system technology, Precision Agriculture, 25(1): 146-185.
https://doi.org/10.1007/s11119-023-10063-3

Al-Kfairy M., Mustafa D., Kshetri N., Insiew M., and Alfandi O., 2024, Ethical challenges and solutions of generative AI: an interdisciplinary perspective, Informatics, 11(3): 58.
https://doi.org/10.3390/informatics11030058

Ampatzidis Y., and Partel V., 2019, UAV-based high throughput phenotyping in citrus utilizing multispectral imaging and artificial intelligence, Remote Sensing, 11(4): 410.
https://doi.org/10.3390/rs11040410

Brookes G., and Smyth S.J., 2024, Risk-appropriate regulations for gene-editing technologies, GM Crops and Food, 15(1): 1-14.
https://doi.org/10.1080/21645698.2023.2293510

Cheng Q., and Wang X., 2024, Machine learning for AI breeding in plants, Genomics Proteomics and Bioinformatics, 2(4): qzae05.
https://doi.org/10.1093/gpbjnl/qzae051

Copland R.R., Hanke S., Rogers A., Mpaltadoros L., Lazarou I., Zeltsi A., Nikolopoulos S., MacDonald T., and Mackenzie I., 2024, The digital platform and its emerging role in decentralized clinical trials, Journal of Medical Internet Research, 26: e47882.
https://doi.org/10.2196/47882

Davison J., and Ammann K., 2017, New GMO regulations for old: determining a new future for EU crop biotechnology, GM Crops and Food, 8(1): 13-34.
https://doi.org/10.1080/21645698.2017.1289305

Ezzelle J., Rodriguez-Chavez I., Darden J., Stirewalt M., Kunwar N., Hitchcock R., Walter T., and D'Souza M.P., 2008, Guidelines on good clinical laboratory practice: Bridging operations between research and clinical research laboratories, Journal of Pharmaceutical and Biomedical Analysis, 46(1): 18-29.
https://doi.org/10.1016/j.jpba.2007.10.010

Fang X.J., Liang Q.X., 2026, Theoretical foundations and practical evolution of breeding CROs: a systematic summary based on 25 years of service experience, Molecular Plant Breeding, 24(1): 318-328.
https://link.cnki.net/doi/10.13271/j.mpb.024.000318

Farooq M., Gao S., Hassan M., Huang Z., Rasheed A., Hearne S., Prasanna B., Li X., and Li H., 2024, Artificial intelligence in plant breeding, Trends in Genetics, 40(10): 891-908.
https://doi.org/10.1016/j.tig.2024.07.001

Fernandez R., Subramaniam P., and Franklin M., 2020, Data market platforms, Proceedings of the VLDB Endowment, 13: 1933-1947.
https://doi.org/10.14778/3407790.3407800

Goktas P., and Grzybowski A., 2025, Shaping the future of healthcare: Ethical clinical challenges and pathways to trustworthy AI, Journal of Clinical Medicine, 14(5): 1605.
https://doi.org/10.3390/jcm14051605

Gumba H., Waichungo J., Lowe B., Mwanzu A., Musyimi R., Thitiri J., Tigoi C., Kamui M., Berkley J., Ngetich R., Kavai S., and Kariuki S., 2018a, Implementing a quality management system using good clinical laboratory practice guidelines at KEMRI-CMR to support medical research, Wellcome Open Research, 3: 137.
https://doi.org/10.12688/wellcomeopenres.14860.2

Han Y., Wang K., Liu Z., Pan S., Zhao X., and Wang S., 2020, Research on hybrid crop breeding information management system based on combining ability analysis, Sustainability, 12(24): 4938.
https://doi.org/10.3390/su12124938

Hu T., Shen P., Zhang Y., Zhang J., Li X., Xia C., Liu P., Lu H., Wu T., and Han Z., 2025, OpenPheno: An open-access user-friendly and smartphone-based software platform for instant plant phenotyping, Plant Methods, 21(1): 76.
https://doi.org/10.1186/s13007-025-01395-4

Kendall G., Bai R., Błażewicz J., De Causmaecker P., Gendreau M., John R., Li J., McCollum B., Pesch E., Qu R., Sabar N., Berghe G., and Yee A., 2016, Good laboratory practice for optimization research, Journal of the Operational Research Society, 67(4): 676-689.
https://doi.org/10.1057/jors.2015.77

Lassoued R., Smyth S.J., Phillips P., and Hesseln H., 2018, Regulatory uncertainty around new breeding techniques, Frontiers in Plant Science, 9: 1291.
https://doi.org/10.3389/fpls.2018.01291

Liang Q.X., and Zhou Y., eds., 2012, GLP Introduction, Sophia Publishing Group Co., Ltd., British Columbia, Canada, pp. 1-100.

Mahmood U., Li X., Fan Y., Chang W., Niu Y., Li J., Qu C., and Lu K., 2022, Multi-omics revolution to promote plant breeding efficiency, Frontiers in Plant Science, 13: 1062952.
https://doi.org/10.3389/fpls.2022.1062952

Menz J., Modrzejewski D., Hartung F., Wilhelm R., and Sprink T., 2020, Genome edited crops touch the market: a view on the global development and regulatory environment, Frontiers in Plant Science, 11: 586027.
https://doi.org/10.3389/fpls.2020.586027

Mu T., Song Q., Liu Y., and Song J., 2025, Initiating the commercialization of genetically modified staple crops in China: domestic biotechnological advancements regulatory milestones and governance frameworks, GM Crops and Food, 16(1): 450-481.
https://doi.org/10.1080/21645698.2025.2520664

Panwar D., Reddy B., Harini A., Mohapatra R., Giri D., Karthickraja A., and Kumar M., 2025, Emerging technologies in precision breeding for sustainable agriculture: a review, Journal of Advances in Biology and Biotechnology, 28(4): 666-680.
https://doi.org/10.9734/jabb/2025/v28i42226

Purnhagen K., and Wesseler J., 2020, EU regulation of new plant breeding technologies and their possible economic implications for the EU and beyond, Applied Economic Perspectives and Policy, 43(4): 1621-1637.
https://doi.org/10.1002/aepp.13084

Qaim M., 2020, Role of new plant breeding technologies for food security and sustainable agricultural development, Applied Economic Perspectives and Policy, 42(2): 129-150.
https://doi.org/10.1002/aepp.13044

Razzaq A., Kaur P., Akhter N., Wani S., H., and Saleem F., 2021, Next-generation breeding strategies for climate-ready crops, Frontiers in Plant Science, 12: 620420.
https://doi.org/10.3389/fpls.2021.620420

Sempéré G., Pétel A., Rouard M., Frouin J., Hueber Y., De Bellis F., and Larmande P., 2019, Gigwa v2-extended and improved genotype investigator, GigaScience, 8(5): giz051.
https://doi.org/10.1093/gigascience/giz051

Shahriar S., Allana S., Hazratifard S., and Dara R., 2023, A survey of privacy risks and mitigation strategies in the artificial intelligence life cycle, IEEE Access, 11: 61829-61854.
https://doi.org/10.1109/ACCESS.2023.3287195

Smulders M., Van De Wiel C., and Lotz L., 2021, The use of intellectual property systems in plant breeding for ensuring deployment of good agricultural practices, Agronomy, 11(6): 1163.
https://doi.org/10.3390/agronomy11061163

Sumathi D., 2025, AI powered business intelligence platform for real time insight and decision support, International Journal of Scientific Research in Engineering and Management, 9(1): 1-7.
https://doi.org/10.55041/ijsrem46479

Tekic A., Willoughby K., and Füller J., 2023, Different settings different terms and conditions: the impact of intellectual property arrangements on co-creation project performance, Journal of Product Innovation Management, 40(3): 679-704.
https://doi.org/10.1111/jpim.12668

Turnbull C., Lillemo M., and Hvoslef-Eide T., 2021, Global regulation of genetically modified crops amid the gene edited crop boom-a review, Frontiers in Plant Science, 12: 630396.
https://doi.org/10.3389/fpls.2021.630396

Van Etten J., De Sousa K., Cairns J., Dell'Acqua M., Fadda C., Guereña D., van Heerwaarden J., Assefa T., Manners R., Müller A., Pè M., E., Polar V., Ramirez-Villegas J., Solberg S., Teeken B., and Tufan H., 2023, Data-driven approaches can harness crop diversity to address heterogeneous needs for breeding products, Proceedings of the National Academy of Sciences of the United States of America, 120(14): e2205771120.
https://doi.org/10.1073/pnas.2205771120

Varshney R.K., Singh V.K., Hickey J.M., Xun X., Marshall D.F., Wang J., Edwards D., and Ribaut J.M., 2016, Analytical and decision support tools for genomics-assisted breeding, Trends in Plant Science, 21(4): 354-363.
https://doi.org/10.1016/j.tplants.2015.10.018

Xu Y., Zhang X., Li H., Zheng H., Zhang J., Olsen M., S., Varshney R., K., Prasanna B., M., and Qian Q., 2022, Smart breeding driven by big data artificial intelligence and integrated genomic-enviromic prediction, Molecular Plant, 15(11): 1664-1695.
https://doi.org/10.1016/j.molp.2022.09.001

Yan J., and Wang X.F., 2022, Machine learning bridges omics sciences and plant breeding, Trends in Plant Science, 28(1): 199-210.
https://doi.org/10.1016/j.tplants.2022.08.018

Zhao X.Y., Pan S.H., Liu Z.Q., Han Y.Y., and Wang K.Y., 2022, Intelligent upgrading of plant breeding: decision support tools in the golden seed breeding cloud platform, Computers and Electronics in Agriculture, 194: 106672.
https://doi.org/10.1016/j.compag.2021.106672

Zhou W., Yan Z.X., and Zhang L.T., 2024, A comparative study of 11 non-linear regression models highlighting autoencoder DBN and SVR enhanced by SHAP importance analysis in soybean branching prediction, Scientific Reports, 14(1): 5905.
https://doi.org/10.1038/s41598-024-55243-x

Zhu W.C., Li W., Zhang H., and Li L., 2024, Big data and artificial intelligence-aided crop breeding: progress and prospects, Journal of Integrative Plant Biology, 67(4): 722-739.
https://doi.org/10.1111/jipb.13791

# Appendix

Appendix A Proposed Framework for Certification Standards of Breeding CRO Platforms

| Certification dimension | Key Indicators | Evaluation Method | Notes/Remarks |
|---|---|---|---|
| Service Capability | Timeliness of trial execution; reproducibility verification pass rate; client satisfaction | Quantitative scoring (0–5) | Can be evaluated by third-party assessors or via client feedback |
| Data Quality | Data completeness; metadata richness; missing data control mechanisms; error rates | Automated validation+ expert review | Supported by automated quality assessment tools |
| Degree of standardization | SOP documentation coverage; version control mechanisms; deviation records and corrective actions | Document audit+system-generated comparison reports | Relies on evaluation of standardized SOP management systems |
| Compliance Capability | GLP compliance level; completeness of audit trails; data access control management; incidence of compliance events | Checklist inspection+ expert review | Includes GMO and data compliance requirements |
| Level of Intelligence | Number of AI tools embedded in workflows; model performance (AUC/accuracy); availability of explainability tools | Model evaluation+system functionality review | Requires submission of real-world application cases and model documentation |
| Innovation Capacity | Frequency of adoption of new technologies; participation in open-source initiatives; outputs in publications/patents | Expert scoring +literature/patent review | Reflects platform-driven research and innovation capacity |

Notes: It is recommended that this framework be developed with reference to international certification and standards systems such as ISO/IEC 17025, OECD GLP, and FAIR data principles, and progressively evolve into a hybrid mechanism combining industry self-certification and third-party independent auditing.

Appendix B Recommended Items for Breeding Service Platform Regulatory Sandbox Mechanism

| Pilot Theme | Core Testing Content | Expected Policy Output and Goals |
|---|---|---|
| AI-based Decision Tools and Algorithm Traceability | Scope of model use, clarification of attribution and liability, application boundaries of evaluation materials and AI | Determine whether AI-based evaluations, risk assessments, and test recommendations can be used for submission materials |
| Data Cross-border Flow Testing | Multi-location data sharing, crop data cross-border transfer, data security flow | Develop classification management and approval rules for data crossing borders |
| SOP and Electronic Record Compliance | Metadata, signatures, and traceability under the electronic recording framework | Assess whether electronic records comply with data integrity reporting effectiveness |
| Contract Template Compatibility Testing Mechanism | Agreement signing process, version consistency, data authorization and sharing | Promote cross-institution agreement templates for data, IP, and MTA models |
| AI Training and Data Labeling Risk Review Mechanism | Type of datasets used for training, data labeling quality and audit process, and sensitive information filtering | Establish an AI sandbox framework for “Trustworthy AI” governance |

Note: The regulatory sandbox is recommended to be led by the Ministry of Agriculture and Rural Affairs or by local pilot zones or Free Trade Zones in conjunction with the Ministry of Science and Technology's regulatory divisions. Reference can be made to best practices from the financial technology, medical AI, and other sectors

Appendix C AI Governance Evaluation Framework (Applicable to Breeding CROs)

| Governance Dimension | Indicators or Tools | Evaluation Notes |
|---|---|---|
| Model Performance | AUC, RMSE, Accuracy, Precision, Recall | Evaluated based on context; model performance metrics should match the complexity and scale of application scenarios |
| Robustness | Multi-environment data, Performance across environments | Assesses stability under different conditions and ecological scenarios |
| Explainability | SHAP values, Feature importance ranking | Evaluates whether outputs are interpretable and understandable by non-AI experts |
| Fairness and Inclusiveness | Coverage of underrepresented varieties, representation of marginal traits | Evaluates whether model overlooks rare traits or species, or reinforces biased decisions |
| Compliance and Data Privacy | Existence of user agreements/training, differential privacy, access control | Ensures data security and ownership compliance in sensitive contexts like genetic and farmer data |
| Traceability and Reproducibility | Model versioning, training records, task accountability | Assesses whether model development process and outcomes are fully traceable and reproducible |

Notes: It is recommended that this type of indicator system be used as an evaluation reference for breeding platforms participating in national projects, international collaborations, or industry fund-supported projects—thus promoting the transition of AI governance from corporate self-discipline to regulatory coordination