Molecular Plant Breeding 2026, Vol.17, No.1, 1-14
http://genbreedpublisher.com/index.php/mpb

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: 11 Dec., 2025
Accepted: 10 Jan., 2026
Published: 16 Jan., 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): 1-14 (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 toward platform-based collaboration. Its core characteristics are reflected in the synergistic evolution of molecularization, intelligence, and systemization. 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 the development of new varieties. Molecular marker technologies represented by SNPs, together with 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 highly efficient tools for the targeted modification of functional genes, markedly 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 phenotypic data acquisition to become large-scale and real-time, driving breeding into an intelligent stage under the impetus of artificial intelligence and big data (Zhu et al., 2024).

To address the increasing complexity of R&D processes as well as the growing pressures of quality control and regulatory compliance, the CRO (Contract research organization) model has begun to be rapidly adopted in the field of bio-breeding. Originating from the pharmaceutical sector, CROs are a form of specialized service organization that enhance R&D efficiency and reduce compliance risks for innovation entities through standardized workflows, professional teams, and auditable systems. Existing studies indicate that breeding-oriented CROs can effectively address inefficiencies caused by fragmented resources, non-standardized processes, and regulatory pressures in traditional breeding systems, and have become an important pillar of modern breeding systems (Fang and Liang, 2026).

Building upon previous systematic analyses of the evolutionary pathways and platform models of breeding CROs, this study further focuses on the construction of a future-oriented service platform architecture (Fang and Liang, 2026). At present, while breeding CRO platforms are developing rapidly, they also expose structural challenges such as the lack of unified service standards, insufficient regulatory adaptability, and unclear pathways for digital transformation. These issues urgently require breakthroughs in institutional design and system-level capabilities.

This study focuses on three key future directions for breeding CROs-standardization, compliance, and intelligence-and aims to systematically propose a theoretical framework and practical pathways for platform-based CRO services.

Standardization: By establishing unified SOP and GLP systems applicable across species and scenarios, service consistency and data reproducibility can be enhanced (Liang and Zhou, 2012); Compliance: By strengthening regulatory adaptation and risk control capabilities in areas such as transgenic technologies, gene editing, biosafety, and data governance; Intelligence: By integrating 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 seeks to provide a replicable structural solution for breeding CRO platforms to evolve from a “technology execution-oriented” model toward a “system empowerment-oriented” model, and to offer theoretical support and applied demonstrations for quality governance, resource allocation, and the translation of science and technology in modern seed industries.

1 Diagnosis of Industry Status and Challenges

1.1 Lack of standardization impedes mutual recognition of services

The breeding CRO industry is currently in a developmental stage characterized by “multiple isolated initiatives and fragmented governance.” Most institutions lack unified standards in experimental workflows, data collection, field trial management, and quality control systems. Significant disparities exist in GLP (Good laboratory practice) and SOP (Standard operating procedure) frameworks, making service outputs difficult to compare, reuse, and reproduce across organizations (Liang and Zhou, 2012; Van Etten et al., 2023). This fragmentation not only constrains the professionalization of CROs but also hinders their integration into domestic and international regulatory and accreditation systems, thereby affecting high-end clients’ confidence in data reliability (Lassoued et al., 2018; Menz et al., 2020).

By contrast, institutions such as the United States Department of Agriculture (USDA) and the European Food Safety Authority (EFSA) have implemented explicit quality standards and regulatory interfaces for outsourced agricultural R&D services. CRO services in Europe and North America commonly adopt systems such as ISO/IEC 17025 and OECD GLP, enabling cross-border data recognition. In comparison, China still lacks an industry-wide standard framework tailored to breeding services, and this absence of standards has become a major bottleneck constraining the upgrading of breeding CROs.

1.2 Weak compliance systems generate high risks

At present, most breeding CROs have not established systematic compliance management mechanisms covering the entire lifecycle of “pre-experiment-in-experiment-post-experiment” activities. Compliance risks are particularly prominent in areas involving genetically modified organisms (GMOs), gene-edited materials, biosafety, and material transfer. Standardized procedures aligned with regulatory authorities (e.g., the Ministry of Agriculture and Rural Affairs, USDA, EFSA) are often lacking, as are enforceable systems for NDAs, MTAs, and intellectual property allocation (Purnhagen and Wesseler, 2020; Qaim, 2020).

Compliance challenges are further amplified in international collaborative projects. Globally, regulatory approaches to new breeding technologies differ substantially: the United States tends to adopt a product-based regulatory approach, whereas the European Union emphasizes process-based regulation and applies particularly stringent oversight to GMOs (Davison and Ammann, 2017). Such regulatory divergence makes experimental data difficult to mutually recognize across regions and increases the regulatory interpretation costs and operational complexity of CRO services (Menz et al., 2020; Qaim, 2020). Therefore, establishing compliance systems aligned with both domestic and international regulations is a critical prerequisite for breeding CROs to sustainably participate in global collaborations.

1.3 Data silos constrain 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" as the core source of competitiveness. In practice, however, molecular assay data, phenotypic data, trial management records, and compliance documents are typically dispersed across disparate systems, with inconsistent standards and non-uniform formats, preventing cross-platform sharing and reuse (Fernandez et al., 2020; Mahmood et al., 2022).

Due to the lack of unified data interfaces and platform architectures, most CROs are unable to support continuous iteration of machine learning models or conduct large-scale, multi-year, and multi-crop data training. This limitation directly hinders the practical deployment of intelligent functions such as AI-assisted breeding design, trait prediction, and experimental optimization (Yan and Wang, 2022; Van Etten et al., 2023).

Leading international organizations have begun to develop integrated platforms that unify data acquisition, quality control, compliance documentation, and client interfaces. For example, the CGIAR in the United States and the EU's EJP Soil program are promoting interoperability among agricultural data platforms to enhance data openness and reuse value. These experiences demonstrate that building open, standardized, and intelligent data infrastructures is a key direction for the digital transformation of future breeding CROs.

1.4 Ambiguous terminology and regulatory boundaries undermine industry recognition

At present, there is no globally consistent definition of "breeding CRO," making it difficult to clearly distinguish it from general technical outsourcing services, public breeding platforms, and trial contracting organizations. Both academia and industry have yet to reach consensus on critical 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 U.S. Environmental Protection Agency (EPA) and USDA have clarified service qualification requirements, data usage rules, and reporting standards for outsourced services in areas such as pesticides and genetically modified crops. However, in the field of bio-breeding, the role of CROs has not yet been systematically incorporated into regulatory frameworks due to the immaturity of emerging technologies and evolving regulations. As a result, breeding CROs are often overlooked in terms of policy support, accreditation, and public funding, weakening client recognition 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 industry's transition from fragmented service provision to a platform-based industrial model.

2 Platform Construction Recommendations: Reshaping the Future-Oriented Capability System of Breeding CROs

2.1 Core platform architecture: a problem-oriented "three-in-one" capability framework

Challenges and Issues: Existing breeding CRO platforms commonly suffer from fragmented capabilities and disconnected workflows. Standardization, compliance, and intelligence are often developed in isolation, lacking systemic integration. As a result, platforms struggle to support multi-project parallel operations and cross-regional collaboration.

Development Pathway: Future-oriented breeding CRO platforms should establish, at the top-level architectural design, a "three-in-one" capability framework consisting of standardized service processes, full-lifecycle compliance management, and intelligent system integration (Figure 1). Within this framework, the three capabilities should co-evolve synergistically within a single platform, rather than being implemented as linear or additive components (Ezzelle et al., 2008; Smulders et al., 2021; Xu et al., 2022).

This framework emphasizes three key principles:

(1) Standardization as the foundation, addressing the problem of non-reusable processes;

(2) Compliance as the boundary, addressing the problem of non-auditable outcomes;

(3) Intelligence as the amplifier, addressing the problem of data that cannot be transformed into decision-making insights.

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

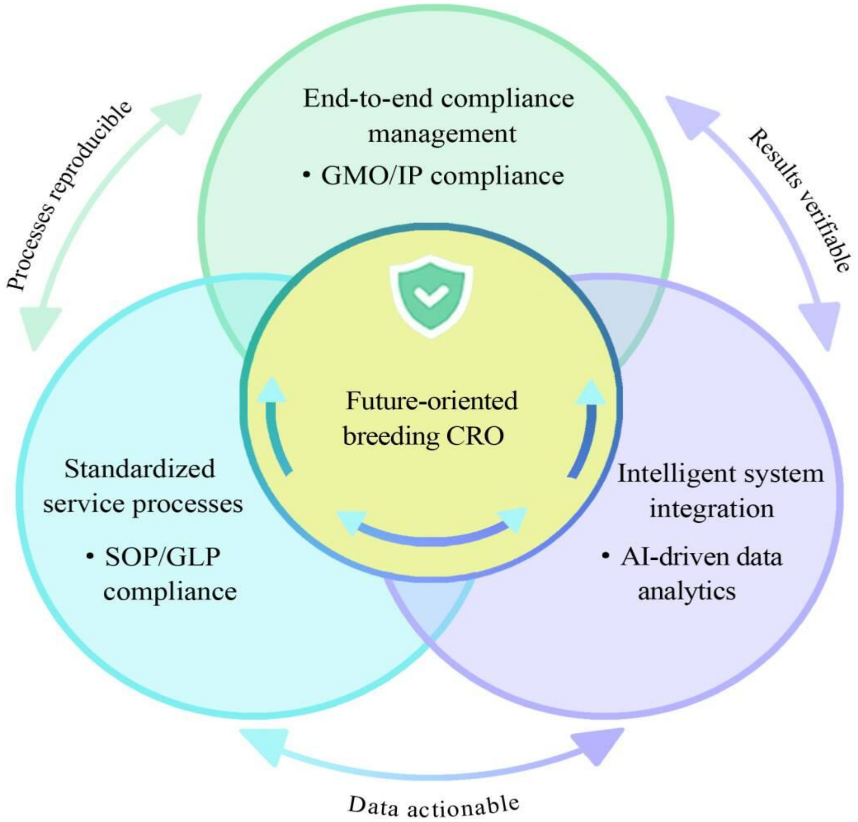

Figure 1 A triadic structural model of the breeding CRO capability system

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

2.2 Building standardization capacity: from fragmented processes to replicable service modules

Challenges and Issues: Although many breeding CROs have established SOP or GLP systems, these are often project-specific and customized, making them difficult to reuse across different crops, teams, and regions. As a result, standardization has not been effectively translated into scalable capabilities.

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 workflow (Figure 2).

At the platform level, SOPs should be abstracted and designed along the following dimensions, rather than repeatedly listing detailed operational steps: Breeding stage dimension (germplasm creation, population development, selection and evaluation, regional trials); Experimental type dimension (molecular assays, phenotypic evaluations, multi-location trials); Risk level dimension (routine experiments, controlled experiments, biosafety-related experiments).

Through an iterative mechanism of “pilot implementation-evaluation-version updating,” SOPs can be transformed from one-off specifications into a “living document system.” When linked with the quality management system, this approach enables closed-loop management of execution, deviation, and continuous improvement (Kendall et al., 2016; Gumba et al., 2018).

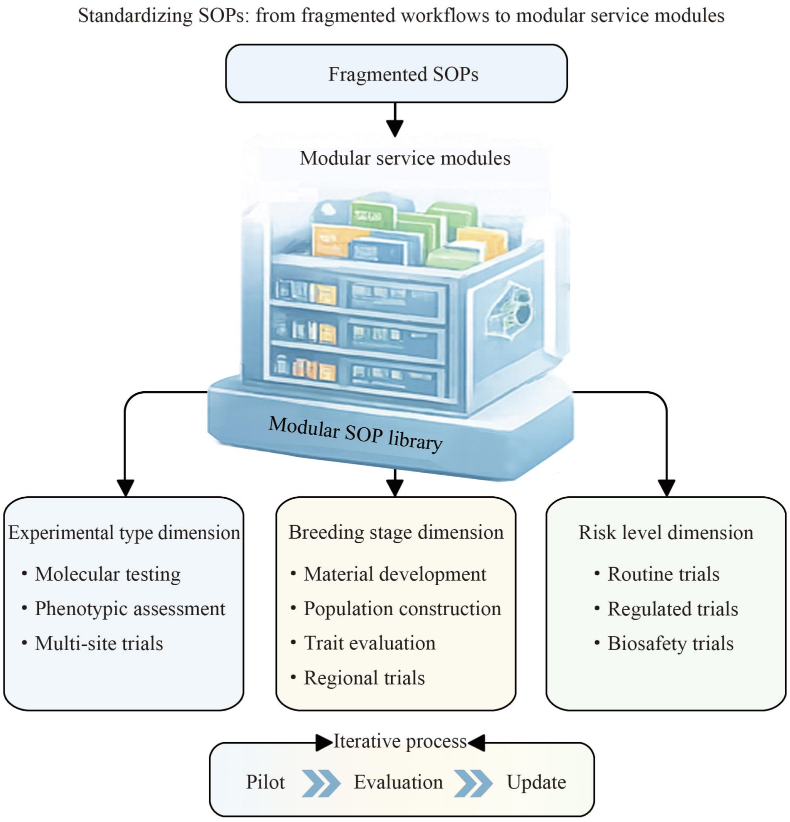

Figure 2 Modular pathway for standardized capability development in breeding CROs
Figure caption: This figure illustrates the transformation of breeding CRO standardization from fragmented, project-specific SOPs to a modular and reusable service framework. By consolidating individual SOPs into a modular 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

2.3 Building compliance capacity: from passive response to embedded governance

Challenges and Issues: Current compliance systems in breeding CROs are largely concentrated on GMO-related projects and rely heavily on manual, experience-based judgment. Such approaches are insufficient to address the systemic risks arising from cross-jurisdictional collaboration, cross-border data flows, and complex intellectual property arrangements.

Development Pathway: Compliance capacity should be upgraded from an “external requirement” to an “endogenous platform mechanism.” By embedding institutional rules, operational processes, and technological tools in a coordinated manner, breeding CROs can achieve replicable and auditable compliance outputs.

At the level of institutional design, it is recommended that platforms adopt a dual-layer structure consisting of a general compliance backbone and scenario-specific adaptation pathways:

General backbone: material source verification-risk classification-trial approval-environmental monitoring-result archiving; Scenario-specific adaptation: loading differentiated regulatory requirements according to target markets (China, the United States, the European Union) (Turnbull et al., 2021; Mu et al., 2025).

At the execution level, NDAs, MTAs, and IP clauses should be bound to specific experimental workflow nodes. Through digital systems, permission controls, audit trails, and document generation can be automatically triggered, thereby reducing uncertainty arising from human operations (Figure 3) (Tekic et al., 2023).

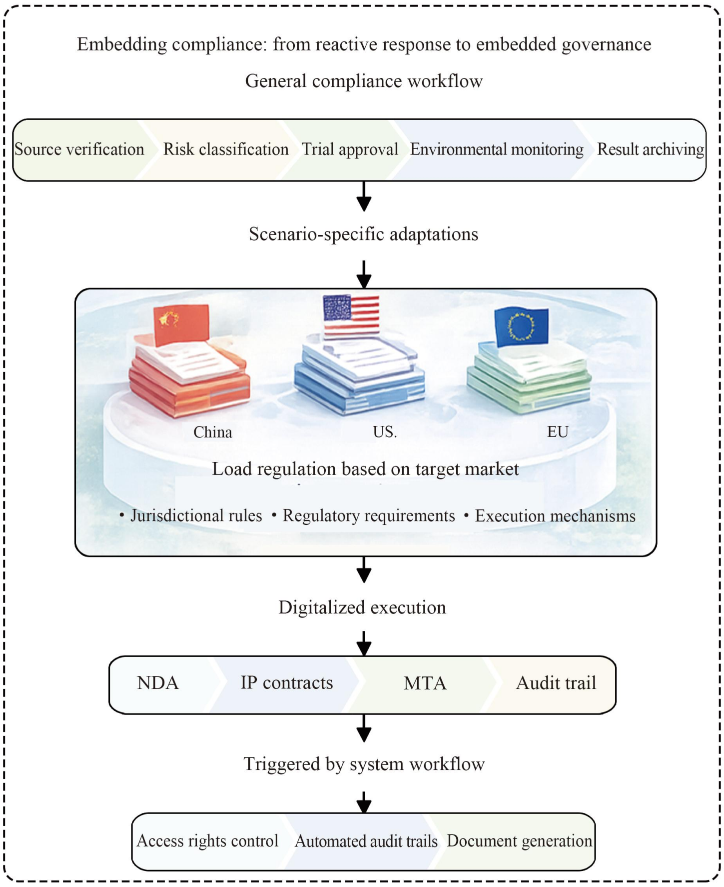

Figure 3 Embedding compliance: from reactive response to embedded governance

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 Building intelligent capabilities: from data accumulation to decision empowerment

Challenges and Issues: Although breeding CROs have accumulated large volumes of molecular, phenotypic, and environmental data, these data are often fragmented in structure and difficult to reuse across projects. As a result, intelligent applications frequently remain at the level of “tool adoption” rather than delivering integrated decision support.

Development Pathway: The development of intelligent capabilities should center on a platform-level data middle layer rather than isolated AI applications.

The platform should prioritize three foundational tasks:

(1) Unifying data models to enable structured integration of genotype-phenotype-environment data;

(2) Introducing AI analysis interfaces for trait prediction, combination optimization, and experimental design recommendations;

(3) Deploying visualization dashboards to translate analytical outputs into interpretable and actionable decision information (Han et al., 2020; Copland et al., 2024).

On this basis, breeding CROs can evolve from “experimental executors” into “intelligent decision-support providers,” significantly enhancing their strategic position within collaborative innovation systems (Sumathi, 2025). As illustrated in Figure 4, intelligent capability development can achieve a transition from data accumulation to intelligent decision-making through the integrated construction of a data middle layer, AI analytical interfaces, and visualization dashboards.

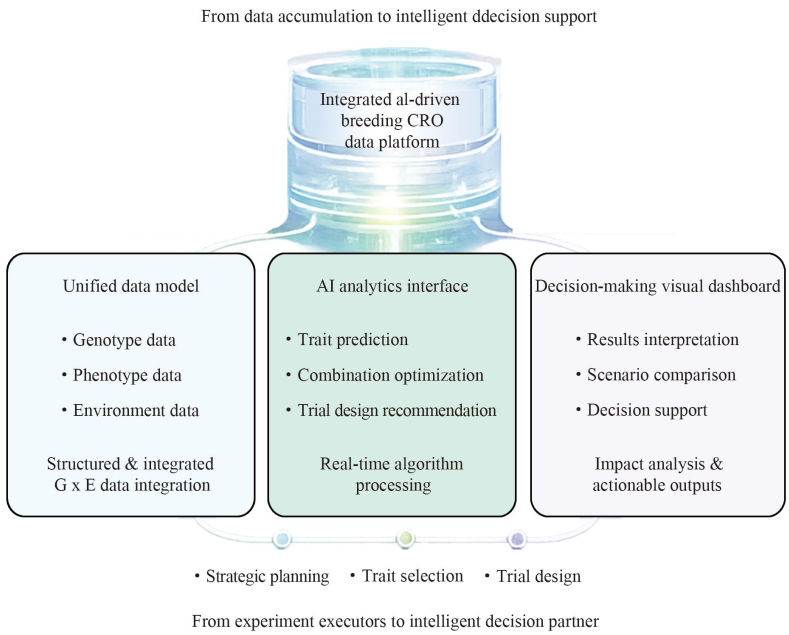

Figure 4 Pathway for developing AI capabilities: from data accumulation to intelligent decision support
Figure caption: Figure 4 illustrates the pathway for developing intelligent capabilities in breeding CRO platforms. It emphasizes the transformation from fragmented data integration, standardization, 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 Platform capability maturity model

To prevent platform construction from remaining at a purely conceptual level, this study proposes a Breeding CRO Capability Maturity Model (Table 1) to guide phased development and 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 model illustrates the evolutionary pathways of platform capabilities across three dimensions-standardization, compliance, and intelligence. The four maturity levels, ranging from Initial to Intelligent, reflect systematic improvements in service workflows, data management, and platform governance. This framework facilitates the assessment of the relative positioning of different breeding CROs in terms of service systematization, digitalization, and internationalization. Overall, the model provides breeding CROs with a clear roadmap addressing the questions of "where they come from, where they are going, and how progress should be evaluated."

2.6 Summary: from capability accumulation to system evolution

By reconstructing platform architecture in a problem-oriented manner and integrating standardization, compliance, and intelligence into a unified capability system, breeding CROs can achieve a transition from "project-based services" to "platform-based infrastructure." This transformation not only addresses current industry fragmentation and compliance pressures but also lays a solid foundation for breeding CROs to assume higher-level roles within the global bio-breeding innovation ecosystem.

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

3.1 Functional evolution of ai in the breeding service value chain

As bio-breeding enters an era driven by multi-omics and data proliferation, the role of artificial intelligence (AI) in breeding services is evolving from a "point-based tool" into a "decision engine," spanning the entire process from sample collection and phenotypic analysis to complex trait modeling and breeding pathway optimization (Xu et al., 2022; Zhu et al., 2024). To enable the true integration of AI into service platforms, technological iteration alone is insufficient; deep coupling with existing information systems-such as LIMS, ELN, and phenotypic recognition systems-is also required to form complete data closed loops and feedback mechanisms (Figure 5).

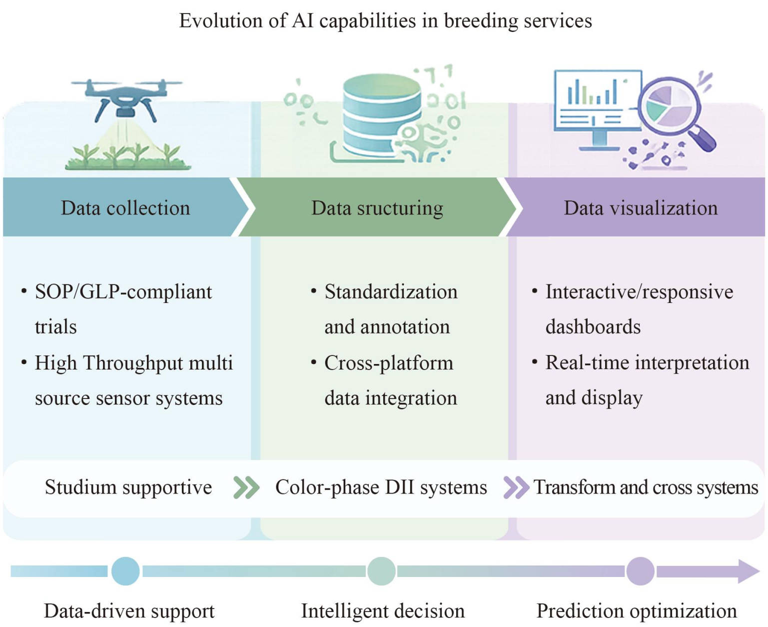

Figure 5 Evolution of ai capabilities in breeding services

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. (This figure was generated with AI assistance)

(1) From "data acquisition" to "structured modeling": integration pathways for intelligent phenotyping systems

In breeding services, phenotypic data collection has long relied on manual operations, characterized by strong subjectivity and poor reproducibility. In recent years, phenotyping platforms based on deep learning combined with sensor networks (such as OpenPheno and PhenoBox) have enabled automated measurements 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 linking them to internal LIMS/ELN systems, a standardized data pipeline encompassing “acquisition-storage-modeling-feedback” can be established.

For example, BASF’s TraitMill platform directly pushes sensor-derived data-such as canopy structure, leaf area, and disease indicators-into an analytical data hub, enabling parallel modeling and differential detection of multiple traits. This integration model based on standardized interfaces allows phenotypic data to be used not only for real-time analysis but also as a stable input source for AI model training.

(2) From “trait prediction” to “multi-objective optimization”: building decision support engines
Complex traits (e.g., yield and stress tolerance) are influenced by polygenic effects and genotype-environment interactions, making them difficult to analyze using traditional methods. AI models-particularly ensemble learning, graph neural networks, and explainable AI approaches-have demonstrated the ability to jointly model high-dimensional traits and perform optimized ranking across multiple objectives (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 lines across diverse environments in real time, providing optimal combination recommendations for each selection cycle. Such systems are often integrated with GIS platforms to incorporate field-level soil, moisture, and pest risk indicators, thereby enabling truly integrated “environment-genotype-phenotype” three-dimensional modeling.

In addition, in maize breeding programs targeting smallholder farming systems, federated learning models have been used to enable shared model training across multiple low-resource countries. This approach has significantly improved selection efficiency for traits related to marginal environment adaptation, demonstrating AI’s adaptability under globally imbalanced data conditions.

(3) From “Breeding Pathway Simulation” to “Strategic Design”: Supporting Forward-Looking Decision-Making
Traditional breeding often lacks mechanisms for dynamically evaluating the long-term impacts of strategic decisions. Contemporary AI systems can now construct in silico breeding platforms, leveraging genomic estimated breeding values (GEBVs), multi-generation simulations, and constraints on genetic diversity maintenance to generate intelligent recommendations for parent selection, mating strategies, and generation advancement (Farooq et al., 2024; Zhu et al., 2024).

These systems can be integrated with front-end experimental design modules of CRO platforms, forming closed-loop linkages with trial databases and phenotypic feedback systems. This enables full-process support for “simulation-execution-correction,” thereby establishing a continuous optimization cycle grounded in real-world feedback.

3.2 Deep integration of platform architecture and AI systems
The development of intelligent platform capabilities is often constrained by the problem of “algorithm silos.” To address this, AI should be embedded into existing management systems, operational workflows, and user interaction interfaces to form a complete service value chain.

(1) Modular design: supporting multi-crop and multi-task heterogeneous configurations
To accommodate diverse application requirements, platforms should adopt a microservices architecture to deploy different AI modules (e.g., phenotypic recognition, trait prediction, simulation and recommendation). Containerized deployment can support flexible invocation across crops and projects (Varshney et al., 2016; Zhao et al., 2022).

For example, the “Golden Seed Cloud” platform decomposes molecular marker analysis, field image processing, and AI recommendation functions into reusable modules. Through a flowchart-style, “building-block” configuration interface, users can rapidly assemble multi-task workflows (Zhu et al., 2024). This model effectively lowers the entry barrier 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 phenotyping platforms are often developed by different vendors, resulting in inconsistent data standards. Breeding CRO platforms should build API frameworks around international standards such as BrAPI and MIAPPE, unifying invocation logic and reducing integration costs (Sempéré et al., 2019).

For instance, OpenPheno achieves seamless connectivity with external phenotyping platforms and data hubs via BrAPI, while the EBS system integrates with enterprise platforms such as SAP and ArcGIS, enabling project management, environmental data, and financial controls to operate within a unified environment.

(3) User-centered collaborative visualization platforms
Algorithmic capabilities only create value when they are effectively used by end users. Platforms should provide visual dashboards combined with collaborative operation interfaces to enable graphical outputs for trial progress tracking, environmental response analysis, and model prediction results (Zhao et al., 2022).

Platforms such as OpenPheno and PhenoApp already offer components including heat maps, timelines, and GIS layer overlays, supporting result sharing and online discussion among project members. This design is particularly suited to breeding project teams operating across multiple locations and roles, enabling cross-regional collaboration while maintaining effective access control and data security.

3.3 Building AI-driven “system-level platform service capabilities”

The development of intelligent platforms should not be regarded merely as a technological upgrade, but rather as a systematic reshaping of the service paradigm of breeding CROs. Such platforms 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 interconnection, model-driven intelligence, and service collaboration to build intelligent hubs with continuous learning capabilities, platform openness, and international adaptability, thereby truly realizing a transformation from “experimental outsourcing” to “co-construction of intelligent breeding ecosystems.”

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 diffusion of digital tools, breeding CROs are gradually transforming from traditional “outsourced service providers” into core nodes within breeding systems. In the future, breeding CROs characterized by platformization, intelligence, and high compliance will become digital infrastructure and collaborative innovation hubs within the global seed innovation ecosystem. Their role will extend beyond experimental execution and data analysis to connecting research institutions, enterprises, regulatory agencies, and international partners, thereby supporting complex breeding projects involving multiple environments and stakeholders worldwide (Xu et al., 2022; Zhu et al., 2024). Within this positioning, CRO platforms must establish end-to-end capability loops encompassing data integration, intelligent decision-making, experimental execution, and compliance support, thereby driving genetic gain while providing 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 in terms of technology and organizational form, their long-term, healthy development remains highly dependent on institutional and standard-based support. The industry currently faces widespread challenges such as fragmented workflows, inconsistent quality standards, and incompatible data interfaces, which severely constrain platform interoperability and service scalability (Brookes and Smyth, 2024; Panwar et al., 2025). Addressing these challenges requires action along two dimensions. First, at the industry level, unified SOP repositories, standardized data structures, and quality control indicator systems should be established to achieve service standardization and process transparency. Second, at the regulatory level, "breeding CRO regulatory sandboxes" should be introduced to clearly define data compliance boundaries, rules for AI tool usage, and platform certification mechanisms, thereby providing controlled environments and institutional safeguards for platform innovation (Alexander et al., 2023; Goktas and Grzybowski, 2025). Only through the coordinated advancement of policy guidance and industry collaboration can breeding CROs form a trustworthy ecosystem characterized by high quality, auditability, and mutual recognition.

4.3 AI-driven capability leap: toward an era of intelligent and 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 increasingly being embedded across all stages of breeding services (Zhou et al., 2024). However, the true 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 workflows. Accordingly, CRO platforms should evolve from "AI tool application" to "AI-driven platform construction," forming intelligent systems in which data and models co-evolve, and algorithms and experiments are tightly coupled. On this basis, comprehensive lifecycle AI governance frameworks-covering algorithm interpretability, model fairness, accountability in decision-making, and ethical boundaries-must also be established to ensure the sustainability and trustworthiness of AI applications (Shahriar et al., 2023; Al-Kfairy et al., 2024).

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

Author Contributions

Xuanjun Fang and Qixue Liang conducted the research, including literature review and data analysis, and were responsible for drafting and revising the manuscript. Both authors have read and approved the final version of the manuscript.

Acknowledgements

The authors sincerely thank the two anonymous peer reviewers for their careful reading of the manuscript and for their valuable comments and suggestions. The authors also thank Ms. Chunyan Tan of the Hainan Provincial Bioengineering Association for providing professional PS image editing support for this paper.

Reference

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

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

Fang X,J., and Liang Q.X., 2026, Theoretical foundations and practical evolution of breedingcros: a systematic summary based on 25 years of service experience, Molecular Plant Breeding, 24(1): 318-328.

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., 2018, 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., 2012, GLP introducion Sophia Publishing Group Inc, 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