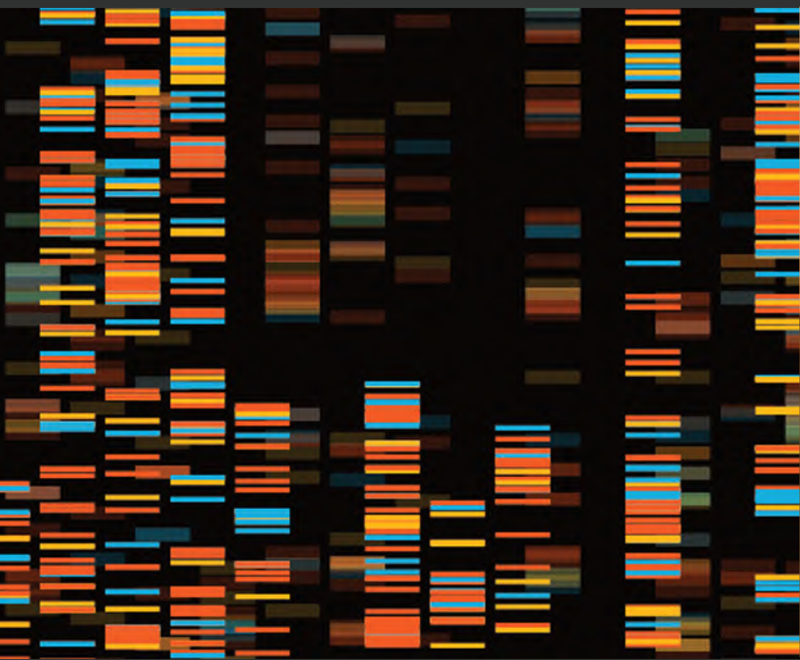
SCIENCE AND INDUSTRY ENDOWMENT FUND

TRACTUUM

SIEF GENOMICS PHASE 2 DIGITAL INITIATIVE IMPACT ASSESSMENT



June 2022

Date	Role	Name
June 2022	Author	Harmeet Kaur, Tractuum
	Final Approval	SIEF Office
	Reviewers	Melissa Straffon SIEF Nerida Gibb, SIEF Thomas Keenan, Tractuum
	Reviewers and Discovery Workshop participants	Alex Whan, CSIRO Brendan Boesen, CSIRO Dadong Wang, CSIRO Nicola Cowen, CSIRO Matt Adcock, CSIRO Radoslaw Suchecki, CSIRO

Copyright and disclaimer

Important disclaimer

Tractuum Pty Ltd advises that the information contained in this publication comprises general statements based on evaluation research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and/or technical advice. To the extent permitted by law, Tractuum Pty Ltd (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

Contents

Figures

Tables

Glossary

Ag&F	Agriculture and Food
AI	Artificial Intelligence
ANU	Australian National University
AR	Augmented Reality
BCR	Benefit Cost Ratio
BD	Business Development
BU	Business Unit
CBA	Cost Benefit Analysis
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CR	Commercial Readiness
DW	Discovery Workshop
EDP	Experimental Development Program
Genomics P2	Genomics Digital Initiative Phase 2
GS	Genomic Selection
ICT	Information and communications technology
IP	Intellectual Property
LCA	Life Cycle Assessment
LIDAR	Laser Imaging Detection and Ranging
MAS	Marker Assisted Selection
ML	Machine Learning
POC	Proof of concept
PS	Producer Surplus
PU	Polyurethane
PV	Present Value
R&D	Research & Development
RD&I	Research, Development, and Innovation
SIEF	Science and Industry Endowment Fund
SNP	Single Nucleotide Polymorphism
SOI	Sequence-of- interest
TBL	Triple Bottom Line
TRL	Technology Readiness Level
UAVs	Unarmed Aerial Vehicles
WGS	Whole Genome Sequencing
WP	Work Packages

1 Executive Summary

Global Challenge -Food Security For more information, see Section 2	There is a need for solutions to feed a global population of 10 billion and meet an expected >50% increase in food demand by 2050.[1] The demand for crop production is expected to grow at 2.4% every year; however, the incremental crop yield is limited to 1.3% and production yields continue to progressively approach their limits. A variety of strategies are being devised to improve crop performance through genetic improvements in novel varieties.[2] However, the current rate of improvement is still insufficient to meet sustainability and food security needs.[3]
Opportunity For more information, see Section 2	The field of molecular biology (MB) offers a great potential to act as a key to addressing some of these challenges. MB unravels the relationship between DNA (the genome), RNA (the transcriptome), and proteins (the proteome) to find the underlying cause for significant crop traits such as disease susceptibility, drought resistance and protein content. As all observable traits have a genetic footprint, selecting for traits based on an understanding of the underlying genome is a possible way of breeding with higher accuracy in a shorter time to deliver improved productivity, resilience, and value generation in breeding programs.
The Industry Challenge For more information, see Section 2	Traditional breeding programs (animals and plants) are long processes as selecting/ breeding for any given trait encompasses multiple life cycles until the observable trait is stably incorporated into the species. Correlations between genomic and phenomic variations can be leveraged more broadly to improve breeding programs. However, the industry faces several challenges associated with the available techniques, such as: • They rely on the acquisition of accurate, diverse phenotypic data, collected across a diverse population, at different stages of individual development. The data collection costs tend to limit some or all of these aspects, thereby reducing the effectiveness of association studies, or the development of robust breeding models. • Many breeding populations, especially those used for introgression, are known to contain significant structural variations, including transpositions, inversions, insertions, deletions, polymorphisms, and repeats. Such variations may potentially impact the expression of traits, yet are fundamentally ignored by marker-based techniques, thereby limiting their power and accuracy. • Few techniques exist which can use transcriptomic or proteomic information and even fewer that incorporate information from two or more 'omes.

[1]www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/HLEF2050_Global_Agriculture.pdf
[2]https://www.mdpi.com/2073-4395/8/12/277/htm
[3]https://www.frontiersin.org/articles/10.3389/fpls.2019.00714/full

The Response For more information, see Section 2	The Science and Industry Endowment Fund (SIEF) supported the Genomics Phase 2 Digital Initiative (Genomics P2) initiative during FY2021-FY2022 under two work packages (WP) to respond to a well-defined need for next-generation tools and platforms that allow the efficient exploitation of genomics-phenomics data to facilitate genomic breeding strategies. These include: - **Pan'omics Toolbox** (WP1) delivered a suite of tools to support the integration and analysis of datasets comprising data from multiple 'omic layers and individuals. The work focused on the development of an efficient method to capture micro and macro-scale commonality and variation within genomic data. The tools were validated using real-world, gigabase-scale genomic (DNA) and transcriptomic (RNA) datasets for species such as wheat and cattle. - **Video Phenomics** (WP2) delivered an Artificial Intelligence (AI)-based prototype phenotyping platform by utilising advanced computer vision and machine learning (ML) methods to improve the automatic collection of crop traits for plant growth monitoring, quantification, and yield prediction applications. The work focused on enabling the efficient and accurate capture of phenomics data for wheat, barley and oats using low-cost sensors mounted on drones and ground vehicles, and through the utilisation of augmented reality (AR). Successful application of outputs has the potential to contribute towards achieving the national target of $100B in farm gate output by 2030.
CSIRO Challenges	Food Security and Quality, Future Industries
Background IP	Video Phenomics builds on the SIEF funded Genomics Phase 1 Digital Initiative undertaken in 2020.[4]
Impact Assessment Approach For more information, see Section 4	The evaluation adopted CSIRO's Impact Framework to identify the logic model of the initiative. <u>The assessment of the prospective impacts of SIEF's investment in the Genomics P2 initiative was conducted based on the adoption of the WP outputs for advancing the plant and animal breeding programs within the Australian agriculture sector through improving conventional breeding schemes and increasing genetic gain.</u> This use case is based on currently explored applications by the project team. The impact hypothesis was developed post-discussion with the core team. Due to limited engagement at this stage, Tractuum did not have the opportunity to consult external stakeholders for the purpose of this assessment. Given that there is no clear/ direct pathway to impact at the current technology readiness levels of either WPs, quantification of prospective impact is not possible at this time. The assessment discusses the prospective benefits from the development and adoption of WP1 and WP2 separately for ease of understanding.
Prospective Impacts	WP1 and WP2 are expected to generate direct benefits from transactions such as the licencing of intellectual property (IP). However, key benefits will be concentrated to sector-level impact through raising the benchmark understanding and awareness of industry (as well as academia), to spark new innovations and design efficient and precise breeding strategies at the industry level. There is synergy between the two packages and their combined capability is projected to have an enhanced impact potential when used together as part of the digital breeding workflow. However, each package has its own impact outside of improved breeding efforts and even within breeding practices, each can be used independently of the other to deliver benefits.

[4] https://sief.org.au/nicta-gift/future-national-ict-industry-platform-program/genomics-digital-initiative/

	Economic	Environmental	Social
	WP1		
	- Research quality, productivity, and efficiency - New services, products, experiences, and market niches - National Economic Performance - New Jobs	- Lower environmental footprint of agricultural activities - Conservation of native plant populations	- Health and wellbeing - Food security - Innovation and human capital (creativity and invention)
	WP2		
For more information, see Section 6 and 7	- Efficiency and Productivity - Research quality, productivity, and efficiency - Management of risk and uncertainty - New services, products, experiences, and market niches - National Economic Performance	- Lower environmental footprint of agricultural activities - Conservation of natural plant populations	- Health and wellbeing - Food security - Innovation and human capital (creativity and invention)

Key impact risks For more information, see Section 6 and 7	**WP1** - Low adoption rates due to slower than anticipated decline in whole genome sequencing (WGS) costs - Failure of R&D hypothesis for application of the SOI Pangenome toolbox as a useful technique for functional analysis and/or genomic-driven breeding **WP2** - Optimal performance of WP2 requires adherence to technical parameters, which is not always met at deployment sites. - End-user site limitations to optimally use the Video Phenomics platform - Lack of applicability across multiple crop systems and platform limitations - Inability to perform timely validation and optimisation of WP2 outputs
Recommendations	- Engage market analysis and associated product development teams to chart a realistic and practicable pathway to commercialisation and the adoption of potential future outputs from these work packages. - Both WP outputs are at a low level of technical maturity. Keeping the needs of key customers/ early adopters at the heart of technical interventions, as well as their early engagement by way of a robust uptake plan from the outset would help focus on shorter-term outcomes while targeting longer-term impacts - Stakeholder engagement and communication as well as a clear and compelling value proposition of the work - Impact thinking and planning from the outset

Support requirements For more information, see Section 8	- Executive sponsorship is required to drive scientific development as well as the proposed (See Section 5.4) bridging activities within CSIRO for uptake of WP outputs for potential applications in other research projects - Involvement in CSIRO projects that offer bridging activities to tie the needs and requirements of these projects with the capabilities of WP1 and WP2 to enhance their research outcomes has the potential to support the testing, validation, and advancement of WP outputs. - With the completion of SIEF funding, both R&D teams require new sources of dedicated funds and resources to pursue the TRL and Commercial Readiness (CR) advancement of the respective WP outputs. - There is a need for cross-disciplinary skills and dedicated resources to advance the TRL and CR of both work packages.
Confidence rating in impact assessment For more information, see Section 10	The Genomics Phase 2 Digital Initiative is aspirational. Although the adoption of WP outputs into the breeding programs are the key focus at this stage, the work has the potential of transferability and value creation across different sectors. The initiative provides a good example of R&D projects that may not produce immediate commercial results but have the potential to push scientific frontiers to drive strategic crop research to deliver against the challenge of Food Security. The technical and commercial viability of technology to deliver any real-world impacts remains unclear at this early phase and at low maturity levels. The projection of benefits is based on several hypotheses and hence the confidence rating in this impact assessment is rated very low by Tractuum. If and as CSIRO's research progresses, the current study should be revisited and refined, drawing on more detailed evidence to provide greater insights to inform decision making.
SIEF's role For more information, see Section 9	SIEF provided critical support to establish proof of concept for commercial use-cases from this research. The funding helped lay the scientific groundwork to develop technologies and software platforms to deliver prospective impacts covered in this report. The support also provided a significant basis to support further scientific and financial interventions as well as collaboration decisions for the future development of the WPs. Future evaluations post the TRL, and CR advancement of R&D must acknowledge attribution of the SIEF investment outputs originating from this initiative as inputs into their research pathways to impact.
Business Units (BUs) and Programs	CSIRO Data61 and Agriculture and Food (Ag&F), CSIRO Future Protein and Drought Missions
Collaborators	Australian National University (ANU), InterGrain
Further Information	SIEF Office; Anne-Maree Dowd, CSIRO

2 Introduction

What is CSIRO’s Genomics initiative?

The Genomics Phase 2 Digital Initiative (Genomics P2) is a CSIRO research effort supported by the Science and Industry Endowment Fund (**SIEF**) to address Australia’s food security and sustainability challenges through advancing genomics and ’omics science. Genomics P2 targets the development and adoption of optimal technologies to catalyse value creation from breeding programs in the form of improved agricultural productivity and resilience. The adoption of the project outputs in the form of practical and novel technologies capable of mapping the static genetic makeup of an individual (the genome) through the dynamic intermediate molecular states (including the transcriptome and proteome) at speed and accuracy, to their observable characteristics (i.e., the phenome), such as disease state, grain yield or fibre quality has potential to accelerate productivity and efficiency gains for Australian agriculture producers. The successful application can contribute towards achieving the national target of $100B in farm gate output by 2030.

Genomics is recognised as a promising field with several potential commercial and clinical applications to contribute to Australia’s economy, society, and the health and well-being of individuals. As at March 2022, the Australian Government’s *Genomics Health Futures* Mission is investing around $500 million in genomic research.[5]

Global Challenge - Food Security

Feeding and clothing the world sustainably are current pressing challenges. There is a need for solutions to feed a global population of 10 billion and meet an expected >50% increase in food demand by 2050.[6] The demand for crop production is expected to grow at 2.4% every year; however, the incremental crop yield is limited to 1.3% and production yields continue to progressively approach their limits. It would require two planets to meet these demands unless there are higher efficiency options that can help achieve this using the currently available land areas, but with much less water, fewer nutrients, and agrochemicals.[7] Climate change, associated with global warming, extreme weather events, resource depletion and increasing incidence of weeds, pests, and pathogens, are strongly influencing major cropping systems.

In this challenging scenario, pertaining to agriculture, a variety of strategies are being devised to improve crop performance through genetic improvements in novel varieties.[8] This includes efficient utilization of existing pools of information for large plant breeding populations through efficient high-throughput technologies, big data management tools, and downstream biotechnology, together with molecular techniques towards automation and digitization of breeding. However, the current rate of improvement is still insufficient to meet sustainability and food security needs.[9]

Opportunity

Traditional breeding programs (both animals and plants) are lengthy processes as selecting/ breeding for any given trait encompasses multiple species life cycles until the observable trait is stably incorporated into the genome. The field of molecular biology (MB) offers great potential to act as a key to addressing some of these challenges. MB unravels the relationship between DNA (the genome), RNA (the transcriptome), and proteins (the proteome) to find the underlying cause for significant crop traits such as disease susceptibility, drought resistance and protein content.

[5] https://www.health.gov.au/initiatives-and-programs/genomics-health-futures-mission
[6] www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/HLEF2050_Global_Agriculture.pdf
[7] https://www.theworldcounts.com/challenges/planet-earth/state-of-the-planet/is-the-world-running-out-of-food/story
[8] https://www.mdpi.com/2073-4395/8/12/277/htm
[9] https://www.frontiersin.org/articles/10.3389/fpls.2019.00714/full

As all observable traits have a genetic footprint, selecting for traits based on an understanding of the underlying genome is a possible way of breeding with higher accuracy in a shorter time. Genomic-driven breeding programs utilize the relationship between a particular trait (phenotype) and the fundamental DNA (genotype) of a population of species.

This information can be leveraged to support new avenues for the genetic improvement of agricultural stock to deliver improved productivity, resilience, and value generation in breeding programs in Australia and globally to address the Food Security challenge.

The concept of the genotype-phenotype relationship is transferable across all biological systems. For example, the physical presentation of a disease/ailments in humans and other species, in most cases can be traced back to variations in the genome. Comparative genomic analysis of a healthy population with that of a diseased population will help in the prediction of individual disease risk, enable early detection of disease, and improve diagnostic classification to better inform individualized treatment.[10] Similarly, deciphering plant-pathogen interactions through the construction of various genome-scale models of pathogens and plants presents an avenue for better understanding of all interactions activated during infection and identification of targets for modification in plants, to confer pathogen-specific resistance.[11] Therefore the work is also relevant to other endeavours such as understanding specific disease susceptibility of an individual or a population and biosecurity.

Industry challenge

What are the key scientific challenges in the pathway of exploitation of MB to address the global Food Security challenge?

There are three main approaches to expediting the traditional breeding process by exploiting the genome to phenome relationships: genomic selection (GS), marker-assisted breeding (MAB), and genetic engineering.

Breeding stock selection within GS programs makes use of predictive models that have been developed using a variety of data, including genomic "markers" (e.g. SNPs), pedigree information, environmental observations, and "breeding value" (which is a kind of phenomic data). MAB uses genomic markers to select breeding stock based on previously-determined correlations between markers and particular traits. Genetic engineering uses techniques that directly incorporate or remove regions of DNA based on knowledge of which DNA regions are responsible for beneficial or detrimental traits.

Correlations between genomic and phenomic variations can be used more broadly, e.g. to improve yield through the detection and treatment of susceptibility to disease or other pathogens. These correlations are determined through large-scale population "association studies", that identify statistically likely associations between genomic marker variations and phenomic variation.

The industry faces several challenges associated with these techniques:

- They rely on the acquisition of accurate, diverse phenotypic data, collected across a diverse population, at different stages of individual development. The data collection costs tend to limit some or all of these aspects, thereby reducing the effectiveness of association studies, or the development of robust breeding models.
- Many breeding populations, especially those used for introgression, are known to contain significant structural variations, including transpositions, inversions, insertions, deletions, polymorphisms, and repeats. Such variations may potentially impact the expression of traits, yet are fundamentally ignored by marker-based techniques, thereby limiting their power and accuracy.
- Few techniques exist which can use transcriptomic or proteomic information and even fewer that incorporate information from two or more 'omes.

Furthermore, the decreasing cost of whole genome sequencing (WGS) is seeing an increase in the number and diversity of fully-sequenced and assembled genomes, and there are no known techniques for performing association studies, or developing predictive breeding models, that can take advantage of the increased fidelity and structural information that this data offers.

CSIRO/ SIEF's response

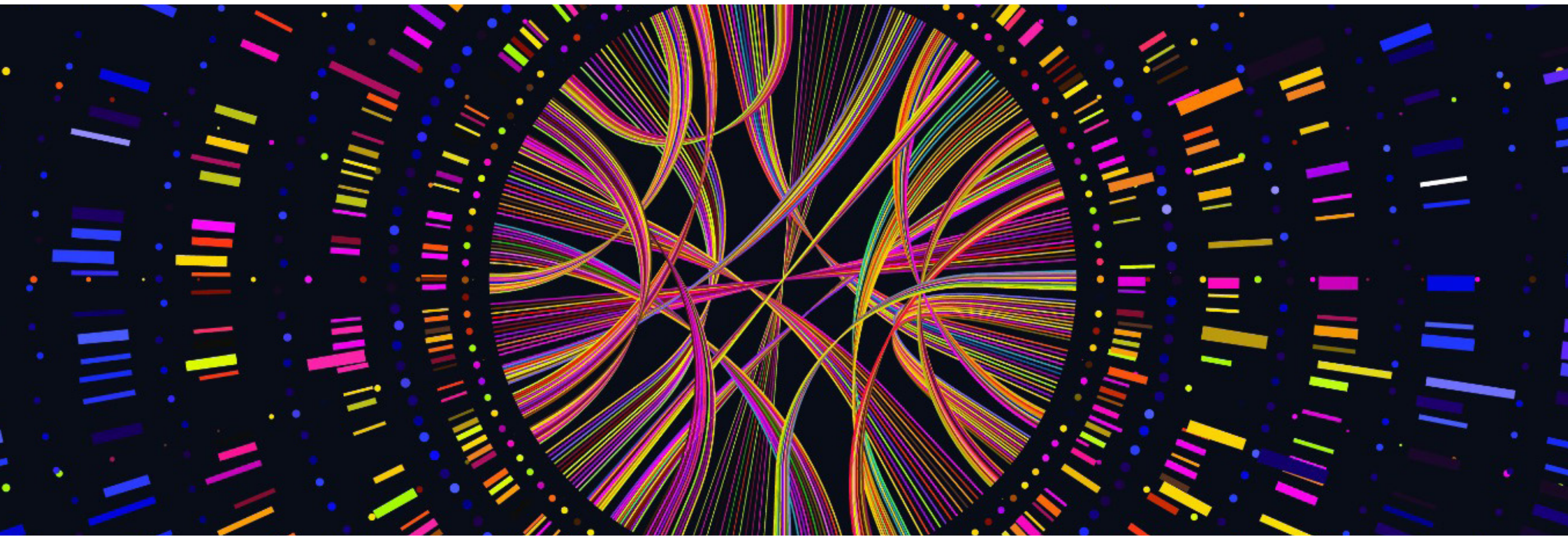
SIEF supported the Genomics P2 initiative to respond to a well-defined need for next-generation, accurate, data science approaches and software prediction

[10] https://www.embopress.org/doi/full/10.1002/emmm.201100153
[11] https://www.frontiersin.org/articles/10.3389/fmicb.2016.01488/full

tools that allow the exploitation of pan'omics data to facilitate genomic breeding strategies and genetic engineering. Under this effort, CSIRO created new 'omics software platforms under two work packages (WP), which include:

Pan'omics Toolbox (WP1) delivered a suite of tools including new pangenome representation and analysis techniques to support the integration and analysis of datasets comprising data from multiple 'omic layers and individuals. The work focused on the development of an efficient method to capture micro and macro-scale commonality and variation within genomic data and was validated using real-world, gigabase-scale genomic (DNA) and transcriptomic (RNA) datasets for species such as wheat and cattle.

This effort aimed at improving both comparative and functional genomic analyses by enabling visualization of the entire genome structure instead of randomly sequenced segments of the genome. The adoption of this computationally efficient approach could help unravel new understandings of the biological cause of significant traits to support genomic-assisted breeding programs and identify new interventions across agriculture, health and the environment.



Video Phenomics (WP2) delivered an Artificial Intelligence (AI)-based prototype phenotyping platform by utilising advanced computer vision and ML methods to improve the automatic collection of crop traits for plant growth monitoring, quantification, and yield prediction applications. This WP advanced the Technology Readiness Level (TRL) of outputs from the Genomics (Phase 1) Digital Initiative to TRL 3-4. See Section 5.4 for further details.

This platform is powered by advanced computer vision and ML methods focused on different genotypes of the three cereal crops, wheat, barley, and oats to enable the accurate capture of phenomics data using low-cost sensors and through the utilisation of augmented reality (AR).

The adoption of the platform has the potential to act as a vehicle to broaden the accessibility and scalability of phenomics technology for commercial breeding applications .

[12]https://sief.org.au/nicta-gift/future-national-ict-industry-platform-program/genomics-digital-initiative/

What is the purpose of Impact Assessment?

The purpose of this ex-ante assessment is to estimate the potential triple bottom line (TBL) impacts from the SIEF Genomics 2 initiative and highlight the critical role played by SIEF support in providing a proof of concept for commercial use-cases from this research. The analysis covers the prospective direct and indirect impacts of the research and development (R&D). The program builds on the SIEF funded Genomics (Phase 1) Digital Initiative undertaken in 2020.[12] The study outlines the impact logic model (Impact Pathway), risks and recommendations, and considers lessons learned during the developmental period to benefit this and other similar initiatives in the future.

To assess benefits from a program of work, a mixed-methods approach is usually adopted to estimate the quantitative (usually by way of Cost Benefit Analysis – CBA) and qualitative benefits. As will be evident from the qualitative discussion in subsequent sections of this assessment, key outputs from the two WPs set the foundations for technology platforms with compelling hypotheses for future impact potential in the form of economic-social-environmental benefits. This includes (but is not limited to) the improved operational and research efficiencies, yield benefits to farmers, creation of time and cost-efficient breeding paradigms, environmental conservation, creation of higher-skilled jobs, collaborative networks, and new technology options created by R&D. However, the technical and commercial viability of technology to deliver any real-world impacts remains unclear at this early phase and low maturity levels. Hence the evaluation highlights prospective qualitative benefits from the program of work.

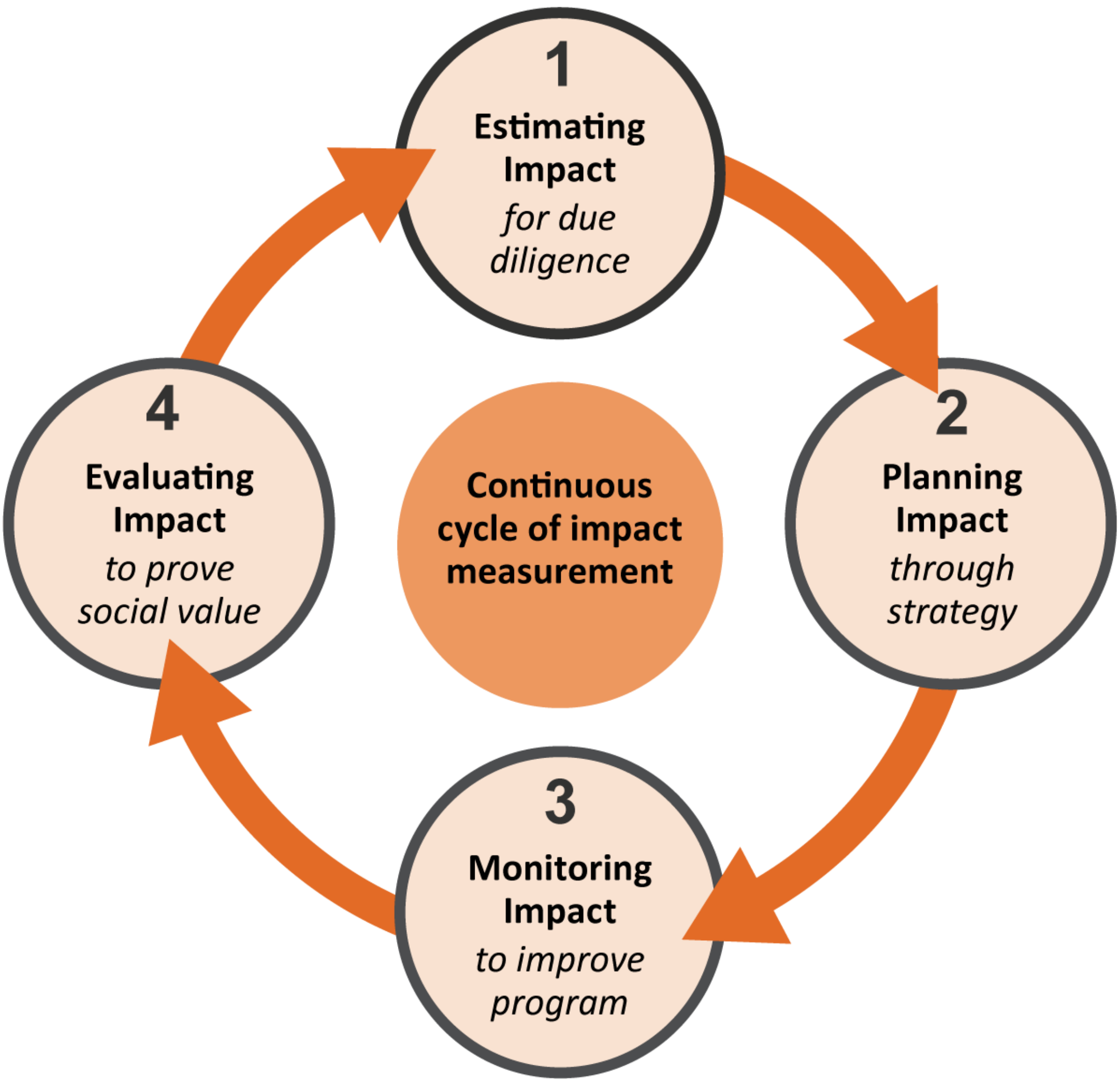


Figure 1: Continuous cycle of Impact measurement objectives[13]

[13] www.hbs.edu/socialenterprise/Documents/MeasuringImpact.pdf

Box 1 Ex-ante Assessment

As will be clear from the assessment, the benefits estimation of the Genomics P2 initiative is largely ex-ante (before realisation) at this stage. An ex-ante evaluation is considered a significant step for impact management of a project throughout its lifecycle as it:

i. provides an opportunity to identify the causal links underlying the investment's path to impact, and allows the investors and others to assess the strengths of these linkages;

ii. helps establish a baseline to effectively determine the changes delivered through the project

iii. promotes due diligence for sound decision-making;

iv. assists with improved implementation, as well as evaluation planning and monitoring

v. allows evidence-based evaluation and reporting later on in the journey of the work.

However, a lack of evidence-based data at this stage makes these assessments highly uncertain.

This report can be read as a stand-alone item, or alongside other CSIRO Agriculture and Food (Ag&F) and Data61 evaluations. The information is provided for the purposes of accountability, communication, engagement, continuous improvement, and future impact management. The audiences include (but are not limited to) CSIRO (especially program and leadership teams, Future Protein and Drought Mission program); SIEF, Australian National University (ANU), InterGrain, Commonwealth, state, and local governments, breeding companies, research institutions, breeders, hardware/software manufacturers, gene banks, digital solution providers and interested members of the public.

3 Background

Journey of Genomics 2 technology development

Background R&D: Genomics (Phase 1) Digital Initiative

In an effort to fill the current lacuna in ’omics data utilisation technologies, CSIRO’s Genomics Phase1 Digital Initiative was started in FY2020 with a focus on the exploration and development of advanced technologies and data science techniques to enhance agricultural productivity and deliver against the challenge of food security. The work was planned to be carried out in three phases.

The Genomics Phase 1 Digital Initiative was segmented into 3 milestones:

1. GENC01 - Immersive data visualisation of population-scale genome architectures.
2. GENCO2: Genome populations as graph data models for genetic trait discovery.
3. GENC03: Utilisation of AI approaches in platforms for real-time data integration and analysis.

The Digital Initiative successfully generated transferable and reusable outputs including immersive visualization tools, graph-based approaches for analysing complex genomics data; foundational research into graph-based representations of population-scale data, and the automation of data collection processes through the development of prototype AI algorithms.

Genomics Phase 2 Digital Initiative

The AI prototype developed during Phase 1 attracted significant industry interest, requiring further R&D to advance the application for commercial exploitation in breeding programs. The foundational graph research from Phase 1 also provided the basis of a potential pathway and a novel software platform for the full exploitation of population-scale genomics, and ’omics more generally (for example proteomics and transcriptomics). This led to the launch of Genomics P2.

Current Scenario

The WP1 team is seeking new collaborations and funding to further refine the capabilities within the Toolbox for comparative genomic analysis, and to develop new capabilities to support functional analysis and genome-driven breeding. There is some internal interest within CSIRO to continue developmental work on the Toolbox, and strong interest from programs focused on particular species (covering livestock, crop, and pest species), to apply the Toolbox to specific research and development problems. There is strong interest from both parties in continuing the project’s collaboration with the ANU, focusing on applying the SOI-based pangenome representation to existing functional analysis and genome-driven breeding techniques. An external partnership with an aquaculture company is in the early stage of negotiations.

The WP2 team is seeking new collaborations and funding to expand the capabilities of the video phenomics platform and including crop traits sensitive to drought stress, such as canopy radiation use efficiency, respiration rates, stomatal conductance, leaf waxiness, canopy temperature, leaf turgor etc, using thermal, hyperspectral and RGB images. This will require addressing issues associated with the effect of data quality on the performance of ML models, upscaling the Video Phenomics platform for large scale crop phenotyping and expanding the platform for other plant species. At the time of this impact assessment, the team had recently submitted a research proposal to apply for CSIRO AI for Mission funding. The proposal is currently under review.

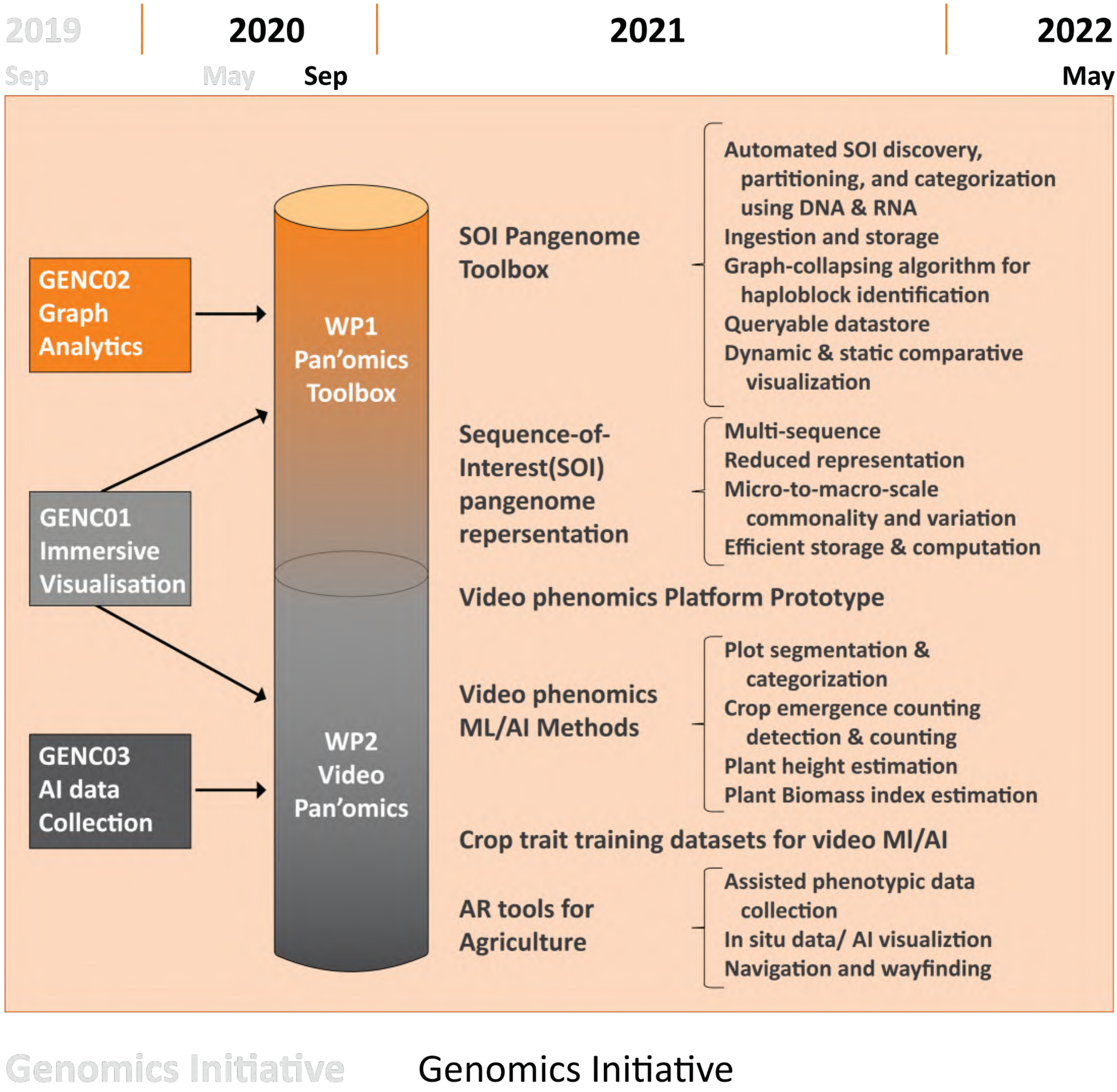


Figure 2: Journey of Genomics Phase 2 Digital Initiative

CSIRO's key challenges in focus

The initiative strongly addresses two of the six CSIRO Challenges, deemed as areas of great significance to Australians in the current environment, namely:

Food Security and Quality: Achieve sustainable regional food security and grow Australia's share of premium AgriFood markets.

Future Industries: Help create Australia's future industries and jobs by collaborating to boost innovation performance.

The work also addresses other CSIRO challenges such as Health and Wellbeing, Resilient and Valuable Environments and A Secure Australia and Region to some extent. CSIRO aspires to create a globally recognised national capability in Genomics by 2030.

IMPACT PATHWAY

TRACTUUM

Impact Statement (Hypothesis): The development and deployment of Pan'omics Toolbox and the Video Phenomics platform to transform the crop and livestock breeding processes and enhance traditional farming practices to increase food security and improve productivity, profitability, and sustainability of Australian crop industries and contribute towards the vision to achieve $100B in farm gate output by 2030.

Enablers (2012-2022)

INPUTS
What we invest

National priority-led opportunity – Enhancing Agricultural productivity

Background IP
- Genomics Phase 1 Digital Initiative

Investment
- CSIRO BU funding and in-kind
- support
- SIEF funding
- InterGrain (WP2)

Capabilities

Specific science skills:
Data- scientists,
Plant breeders
Geneticists
Molecular biologists.

Other
Software Engineering
Communication experts
Business Development
Impact specialists
Change managers

Infrastructure
CSIRO research infrastructure
Ginninderra Experiment Station
CSIRO Boorowa Agricultural Research Station (BARS)

Networks
Partners
Customers / End-users/ Beneficiaries
Supporters

Leadership team support

Regulatory Support

Social Licence to Operate (SLO)

Innovators (2012-2022)

ACTIVITIES
What we do

R&D

- Pan'omics Toolbox (WP1)
- Defining breeders, growers, and scientific researcher uses cases
- Evaluating pan'omic tools and techniques
- Developing novel pangenomic representations
- Developing practical tools for comparative and functional analysis of pangenomic data sets.

- Video Phenomics (WP2)
- Acquiring image and video data for cereal crops at different life cycle stages and across multiple locations
- Collecting ground truth data
- Developing and evaluating AI algorithms
- Defining breeders and growers user cases
- Developing, performance testing and validating AI algorithms and AR prototype

Translation
- Grant applications
- CSIRO internal and industry collaborations

Stakeholder communication and engagement

Mobilise and develop the best talent
- Industry discussion forums and training
- Participation at conferences

Education and Outreach
- PhD and Post-doctoral training

OUTPUTS
Our deliverables

Products

- Pan'omics Toolbox (WP1)
- SOI pangenome representation that captures multi-scale genomic commonality and variation.
- SOI Pangenome Toolbox that provides tools to create, store, query, compare, analyse, and visualise pangenomic data sets.

- Video Phenomics (WP2)
- A web portal as the frontend user interface to upload the images and video data and generate reports
- An AI-based prototype video phenomics platform
- A suite of ML methods for detection and quantification of crop traits.
- Datasets curated for developing ML models for various crop trait extraction.
- Augmented Reality techniques for finding / consuming / creating phenomics data on the farm.
- Virtual Reality Prototyping Environment containing a reconfigurable 1:1 scale replica of the research farm

Capacity and capability building services
- Training programs and consultancy support
- PhD students training
- Cross-disciplinary industry collaborations

Knowledge and data sharing
- Industry/Govt consultations
- New data-sharing networks
- Symposia and conference participation

New data sharing networks
- New collaborations with upstream partners (manufacturers), downstream adopters and intermediaries for TRL advancement and catalysing uptake of solutions
- New networks with interested s takeholders

IP, other publications, and reports
- IP includes: Pan'omics Toolbox, ML models for different trait extraction
- Impact Case Studies
- SIEF technical reports, newsletters
- CSIRO Annual report
- Media articles
- Publications (Journal and Conference)

Scalers (2022 →)

OUTCOMES
The uptake, adoption, or consumption of our work

Pan'omics Toolbox (WP1)	Video Phenomics (WP2)

Adoption

Researchers
- Improved scientific understanding of species evolution through access to better tools (WP1 outputs) to perform genome comparison to enable the phylogenomic analysis to support the discovery of genomic targets.
- Next level of scientific understanding of genome structure at scale and requisite scientific interventions through access to better tools (WP1 outputs) to perform functional and comparative genomic analysis applicable across multiple domains such as agriculture, food security, healthcare, and medicine (non-exhaustive) (WP1)
- Faster interventions from access to farm-wide information in real-time and technology diffusion (WP2).
- Improve the scalability of phenomics technology (WP2)
- New research options

Industry (breeders/ farmers/ medtech companies)
- Improved on-farm collection, curation, and visualisation of trait data (WP2)
- Improved varieties of high yield and resilient crops and livestock species (WP1 and WP2)
- Improved agricultural productivity and agricultural management (WP1 and WP2)
- Uptake of cost and time-efficient, precision breeding strategies (WP1 and WP2)
- ↓ risk and uncertainty for key users

Government/ Community
- Reliable solutions to address the issue of food security through access to higher productivity options
- Farmers and consumers grow and eat better
- Higher quality and healthier food for consumers
- Access to climate-adapted crops for domestic consumption and exports
- Development of capability and capacity in the genomic breeding space

SIEF
- Recognition of the value generated by SIEF on Australia's innovation system through promoting:
  - world-class science
  - fields of research that align with National Science and Research Priorities
  - cross-disciplinary collaborations that deliver solutions to further the interest of Australian communities

CSIRO
- Uptake by CSIRO's internal programs (e.g. Protein Mission and Drought Resilience Mission) to build intellectual property value and help meet mission goals
- Enhanced reputation among breeding and MedTech companies
- Follow-on growth for CSIRO and Australia from the increased competitiveness and reputation in this space
- New research projects as a result of the success of the initiative to generate new avenues of research income as well as opportunities for capability and capacity development

New Collaborations
New networks and collaborations that harness the collective capability of industry, research partners, government, and community to drive advancements in digital agriculture.

New Investments
The establishment of a more coordinated national focus improves key stakeholder traction and improves the inflow of new investments including Government funding programs and private investments to underpin the technology advancement (TRL) and scale-up

Benefits

IMPACTS
Benefits to economy, environment, society

Economic
- Productivity and efficiency
- New services, products, experiences, and market niches
- National Economic Performance
- New Jobs

Environmental
- Lower environmental footprint of agricultural activities
- Conservation of natural plant populations for specific environments

Social
- Health and wellbeing -improvement in living standards
- Access to resources, services and opportunities
- Food security
- Innovation and human capital

Feedback Loop

Figure 3: SIEF's Genomics Phase 2 Digital Initiative - Impact Pathway

5 Impact pathway discussion

6.1 Inputs

Resources applied to deliver activities.

This section provides information on the key inputs invested in the SIEF Genomics 2 initiative.

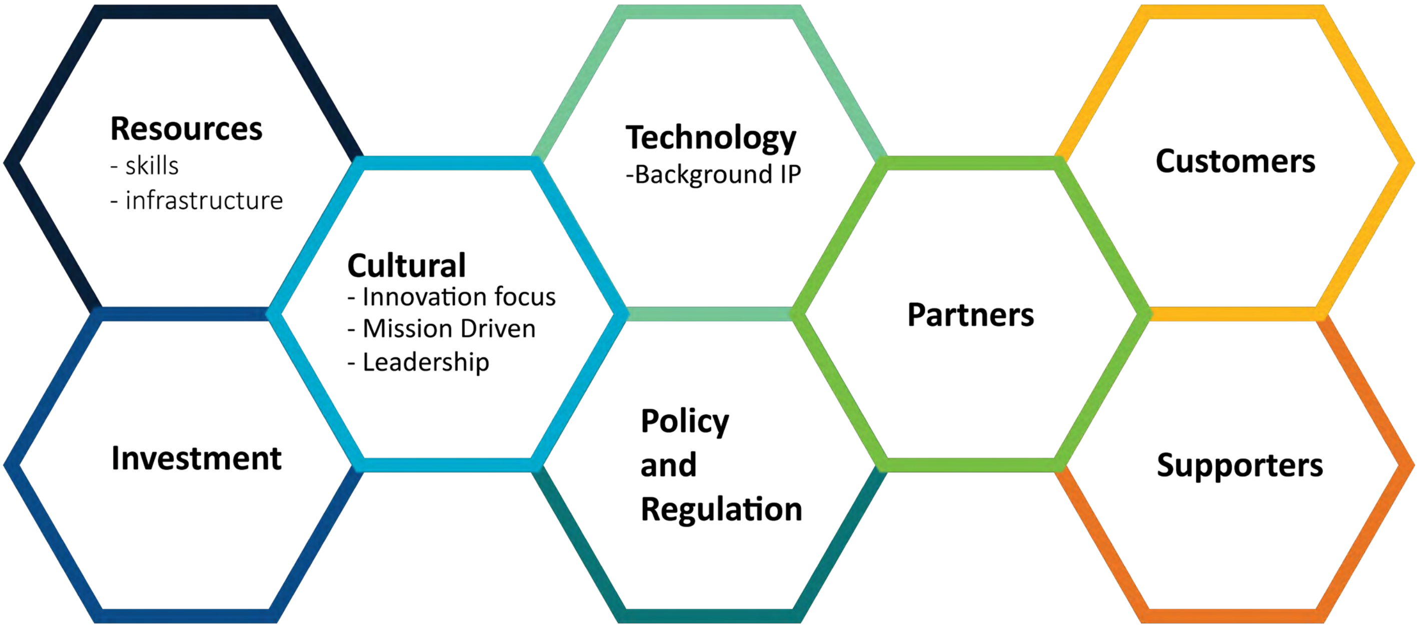


Figure 4: Key inputs to deliver against objectives of Genomics Phase 2 Digital Initiative

The key input requirements to accomplish science deliverables are identified in **Figures 3** and **4.**

Table 1 depicts the investment made by CSIRO, and SIEF for Genomics P2 only. When conducting the quantitative analysis of benefits against costs, outlays associated with Genomics (Phase 1) Digital Initiative and other underlying costs of developing this knowledge (where data is available) should be included in the overall costs as these provide foundational knowledge for this work.

Table 1: Financial and in-kind support for the project

Contributor/type of support	FY2017	FY2017
	Cash	
SIEF		$1,494,897
CSIRO	$3,100,000	
InterGrain	$214,200	$214,200
	In Kind	
CSIRO		

Contributor/type of support	FY2017	FY2017
Overall Annual investment (nominal)	$3,314,200	$1,709,097
Overall Annual investment (real)	$3,430,055	$1,709,097
Overall Annual investment (PV, in FY2022 AUD)	$3,670,159	$1,709,097
Overall investment (PV, in mil FY2022 AUD)	$5.4	

5.2 Activities

Actions taken or work performed through which inputs (such as funds, technical assistance and other types of resources) are mobilised with the intention of achieving specific outputs.

The high-level activities required to accomplish science deliverables are identified in Figure 3. Some additional details for R&D activities are covered below:

Pan'omics Toolbox (WP1)

R&D activities included:

- Conducting a user/stakeholder study to identify use cases for pangenomic and other 'omic data for functional and comparative genomic analyses, as well as for livestock and crop breeding programs.

- Review and analysis of existing pangenome representation techniques and tools, in order to establish their suitability for construction and analysis of large, complex genomes such as wheat.

- Development of a new pan'genomic representation technique, based on sequences-of-Interest (SOIs), designed to support ad-hoc comparative genomic analysis between large, complex genomes in a computationally efficient and informative manner, and with the capability to be used to enhance existing functional analysis and genomic selection processes through the incorporation of information about genomic commonality and variation at multiple scales within a breeding population.

- Development of a suite of tools (the Pan'omics Toolbox) for the construction, storage, comparison, analysis, and visualisation of SOI-based pangenomes.

- Evaluation of the Pan'omics Toolbox' tools and the underlying SOI Pangenome representation technique using real-world genomic datasets.

Video Phenomics (WP2)

R&D activities included:

- Acquiring image and video data for three cereal crops at different life cycle stages across multiple trial locations. Collecting repeat data using sensors mounted on UAV and a mobile ground-based platform, including RGB, depth and thermal videos and laser imaging, detection and ranging (LIDAR) point clouds.

- Collecting ground truth trait data by processing sample plants across the plant life cycle.

- Developing and evaluating AI algorithms for automated detection of different traits. The AI algorithms were developed and assessed from analysing image, video and LIDAR data by comparing the predicted traits from the AI models and the ground truth.

- Evaluating robustness of the AI algorithms using data collected from multiple locations for different crops at different growth stages.

- Conducting user studies involving breeders and growers use cases

- Developing, performance testing and validating AI algorithms and AR prototype

5.3 Outputs

The research solutions, services, and/or capacities that result from the completion of activities within the SIEF Genomics P2 initiative.

Pan'omics Toolbox (WP1)

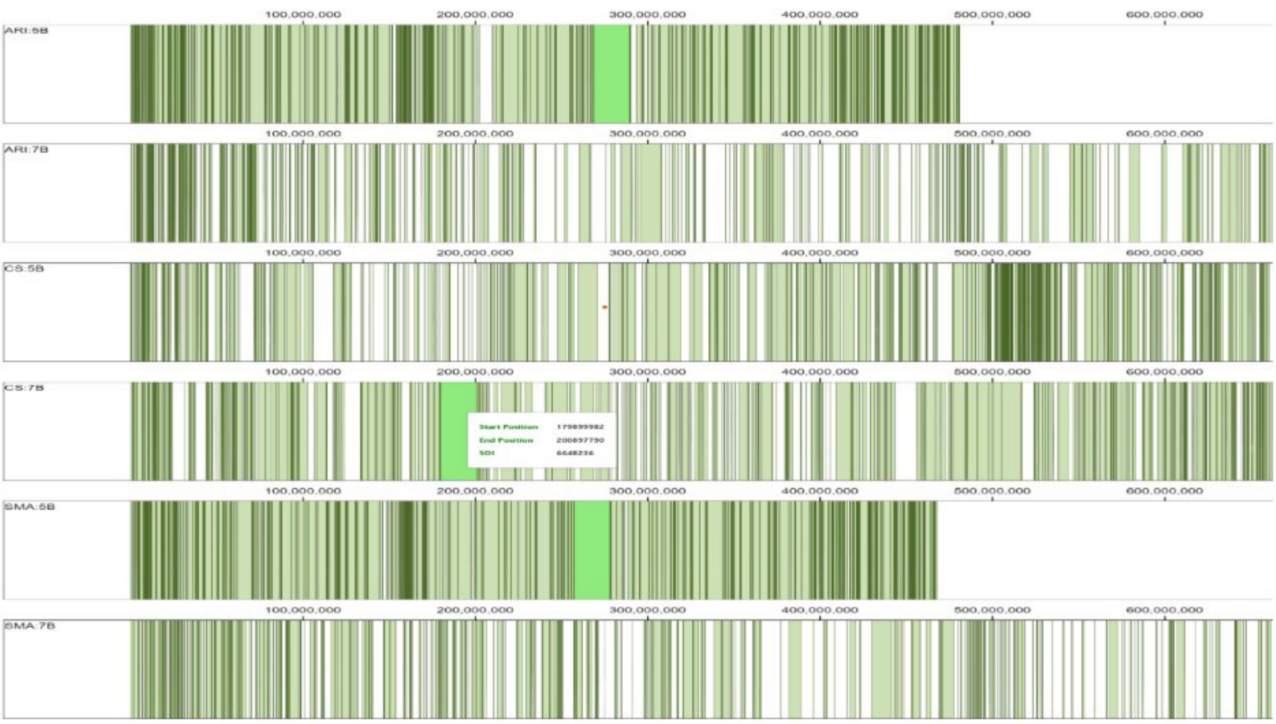
Sequence-of- interest (SOI) Pangenome Toolbox and Sequence-of- interest (SOI) Pangenome representation are key outputs from WP1. See Figure 3.

As a part of this development, the work produced new techniques and technologies for the collection, representation, analysis, and visualisation of pangenomic datasets. The work delivered a suite of tools to allow the SOI Pangenome to be used for practical tasks, and includes facilities to partition genomes into SOIs; categorise SOIs into SOI Types; ingest and store a pangenome; identify larger-scale, multi-SOI commonality and variation; and query and visualise the pangenome.

Key attributes of the outputs include:

- SOI Pangenome Representation
    - Multi-sequence
    - Reduced representation
    - Micro-to-macro scale commonality and variation
    - Efficient storage and computation
- SOI Pangenome Toolbox
    - Automated SOI discovery, partitioning and categorisation using DNA and RNA
    - Ingestion and storage
    - Graph-collapsing algorithm for haploblock identification
    - Queryable datastore
    - Dynamic and static comparative visualisation

The proof of concept for comparative genome analysis has been established. Further development is needed to expand the suitability of outputs to functional genomic analysis. WP1's outputs were validated by creating pangenomes for wheat and cattle using real-world, gigabasescale genomic (DNA) and transcriptomic (RNA) datasets and identifying known genomic targets within one such pangenome that are relevant to real-world breeding programs .



SoiViz representation of (part of)5B and 7B chromosomes. In the dynamic visualisation, corresponding blocks are highlighted(bright green) when hovered over. This make it trivial to quickly spot many large structural varients.

- IP includes: Pan'omics Toolbox, ML models for different trait extraction
- Impact Case Studies
- SIEF technical reports, newsletters
- CSIRO Annual report
- Media articles: '[Boorowa researchers predict crop yields using GoPros]', 23 December 2020
- Publications: Two journal papers under preparation

5.4 Outcomes

The intended or desired medium-term effects /change expected to be realized from the successful uptake of research outputs by users. It usually requires the collective effort of partners (commercialisation partner/ initial user/ end-user etc).

The potential outcomes from the Genomics P2 initiative for the key stakeholder groups are given in Figure 3 above. WP1 is currently at TRL 3 and WP2 stands at TRL 3-4. The pathway to impact for the SIEF Genomics Initiative will require advanced development and uptake of WP1 and WP2 outputs through a combination of cooperative research with other institutions, application to internal CSIRO research and development programs and industry partnerships. Assuming the TRL is advanced successfully, measurable adoption for commercial applications through a suitable commercialisation pathway is a critical ingredient in the translation of the outputs into outcomes and impacts to realise any real-world value.

Industry and academia

The WP1 R&D team is having discussions with industry organisations to build breeding value prediction and enhance sample-based methods. The key focus currently is to forge R&D partnerships for advancing the toolbox's capabilities as well as tailoring and optimising its use for the end-user specific adoption as the next step towards technology adoption and its conversion to real-world benefits. Since the proof of concept for comparative genome analysis has been established, the current focus is on driving the uptake of tools for this application through these industrial collaborations. The group also has an ongoing collaboration with the ANU to pursue the development of WP1 for functional analysis and breeding-value-prediction techniques for multi-sequence pangenomes. As the work progresses, peer-reviewed and published academic papers are expected to attract new research collaborations and technology advancements.

The claimed capabilities of the Video Phenomics platform were also tested with industry partnerInterGrain by comparing the quantified and predicted traits from the tool with the ground truth data from the field. Learnings from this early engagement are being utilized to understand end-user requirements and operational limitations, in order to optimize the platform for improved commercial outcomes.

Box 2 -Targeted application for adoption

The uptake of WPs to augment the breeding systems for future needs is the target application area during the early years of adoption:

WP1 - The adoption of the toolbox to enable population scale, comparative, full-genomic analysis, that forms the basis of activities such as phylogenomic analysis, and discovery of novel genomic targets for breeding.

WP2 – The adoption of the Video Phenomics platform for both traditional (phenotype -driven) and modern (genomics-driven) breeding programs to correlate population-scale phenome and phenotypic variation with an underlying genetic fingerprint. Adoption of the AR arm of the video phenomics toolbox to facilitate AR-guided decision-making and troubleshooting on the ground for farmers and breeders.

5.5 Impacts

Overall impact = Direct impact + Sector level impact

Direct impact: Σ (benefits generated through directly working with customers and partners)

Sector level impact: Σ (benefits generated through raising industry's benchmark understanding and awareness, and sparking new innovations)

WP1 and WP2 under the Genomics P2 initiative are early-stage initiatives of strategic importance due to their potential to facilitate different multidimensional research opportunities and use cases. There are expected to be direct benefits from the adoption of outputs by customers and partners from transactions such as the licencing of IP. However, key benefits will be concentrated as sector-level impact through raising the benchmark understanding and awareness of industry (as well as academia), to spark new innovations and design efficient and precise breeding strategies at the industry level. Given that there is no clear/ direct pathway to impact at the current technology readiness levels of either WPs, quantification of prospective impact is not possible at this time. quantification of prospective impact is not possible at this time.

As briefly noted above, the impact is likely to eventuate first in the Australian agriculture sector

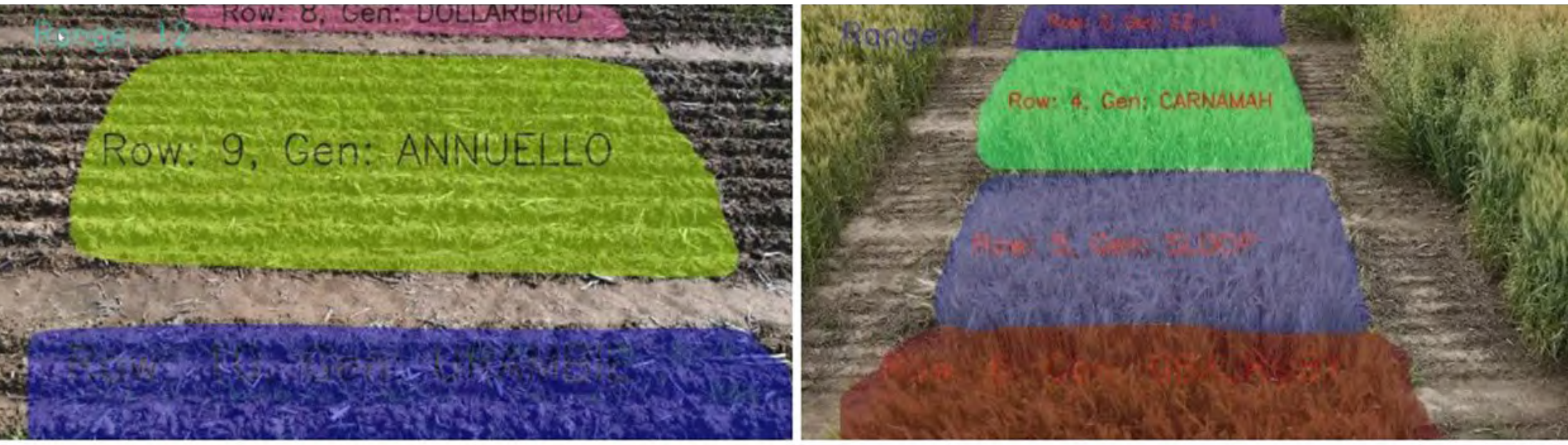
Video Phenomics (WP2)

Video Phenomics platform powered by AI/ML driven detection and quantification, and AR tools for agricultural applications are the key outputs from WP2. See Figure 3.

Key attributes of the Video Phenomics platform output include:

- An easy to access web portal as the frontend user interface that allows users to upload the images and video data and generate reports, using AI-based models to analyse the image and video data to extract crop traits.
- A suite of ML models for detection and quantification of crop traits in wheat, oats, and barley, including:
    - Plot segmentation and categorisation
    - Crop emergence counting
    - Crop spike/spikelet detection and counting
    - Plant height estimation
    - Plant biomass index estimation
    - Curated phenomics research datasets
- AR tools for agricultural applications such as
    - Assisted phenotypic data collection
    - In situ data/ AI visualisation
    - Navigation and wayfinding
    - A simulated research farm in VR

The project utilised images, videos and point cloud data captured at different growth stages of wheat, barley, and oats crops across different farms. The AI based prototype phenotyping platform is capable of automating the detection and quantification of crop traits using farm data collected from equipment such as drones, mounted with sensors like cameras and LIDAR .



Automatically segmented plots at germination(left) and heading growth stages(right)

Capacity and capability building

- Training programs and consultancy support for industry for uptake of technology
- PhD students' training, university collaborations, knowledge sharing through publications (non-confidential information); and
- Cross-disciplinary industry collaborations in this space

Other potential outputs that will propel innovation within CSIRO and more broadly across Australia include:

- New networks of partners, end-users and beneficiaries
- Knowledge and data sharing through collaborative research, industry and government consultations and participation in symposia and conferences (e.g. Wheat Breeding Assembly 2022)

- both plant and animal breeding programs based on discussions with subject matter experts as well as currently explored applications by the project team. A shift in the understanding of the function of gene combinations imparted by the uptake of WP1 and WP2 outputs can help exploit attractive crop traits (e.g. fragrance, texture, etc) much more effectively than previously possible. The improved capabilities have the potential to provide a competitive edge for desirable applications while drawing pathways for accelerated improvement and help produce profitable niche crops some of which were not possible from previous breeding efforts.

Industries associated with human health such as MedTech, pharmaceuticals and precision medicine development are likely to follow (CSIRO per comms) as future end-users for this technology.

Section 6 and 7 discuss the prospective benefits from the development and adoption of WP1 and WP2 as well as the inhibitors in the pathway of delivery of impact. The prospective impacts of the two WPs are considered separately for ease of understanding.

It is important to note that there is synergy between the two packages and their combined capability is projected to have an enhanced impact potential when used together as part of the digital breeding workflow. For example, for a complete digitization and automation of breeding, both work packages can be bridged together so that the pangenome data analysed through WP1 can be correlated with the phenotype data collected through WP2 in both feed-forward and feed-backward modes. However, each package has its own impact outside of improved breeding efforts and even within breeding practices, each can be used independently of the other to deliver benefits.

6 Pan'omics Toolbox (WP1) - Qualitative summary of prospective impacts

Box 3 – Pan'omics Toolbox (WP1) rationale

Background

Current, economically viable genomic breeding strategies (Marker Assisted Selection (MAS) and Genomic Selection (GS)) utilise genomic markers (e.g. Single Nucleotide Polymorphisms (SNPs)) which are only a coarse sampling of the genome. Since these methods are unable to utilise any information about the structure of the genome, they are limited in insight and predictability. Furthermore, these techniques often rely on single reference genomes and are therefore restricted in their capability to take advantage of the full genetic diversity within a species. With the exponential decrease in WGS costs, approaches that take advantage of genomic structure will bring about a new paradigm in genomic analysis and genomic breeding.

Prospective incremental change

The tools developed under WP1 have the potential to provide an otherwise unattainable depth of information with respect to genomic structure with opportunities to find novel insights that cannot be found using current genome sample-based techniques. WP1 leverages full genome information to provide answers to critical questions and lays the foundation for the next generation of tools required to develop products and crops around a specific mechanism of action. This is a key scientific differentiator for tools developed under WP1 compared with other solutions available to end-users.

The adoption of WP1 outputs has the capability to act as an accelerator towards whole genome-based breeding through fundamentally advancing understanding of how a genome manifests into a specific phenotype and the underpinning factors. Because this approach allows tracing the origin of any trait back to the building blocks of life (DNA), its applicability is wide and relevant in all domains dealing with biological processes such as medicine, agriculture, and husbandry.

Key Trends

Current research indicates that WGS based tools are likely to form the basis of next generation of genomics assisted breeding, irrespective of where they are developed. In the interim, it is more likely that hybrid technologies that utilize a mix of both traditional sequencing and WGS will emerge in this space.

For example KeyGene, a European plant research company focusing on technology innovation for crop improvements provides platform-based breeding solutions and, has made considerable advancements in utilising the Pangenomic space.

There is a growing commercial interest in Pan'omics as access to detailed population level whole genome data has wide applications in different fields working with any biological population such as public health, precision medicine to develop genetically efficient cellular machinery towards bioprocess applications, etc.

Prospective Impacts

The potential TBL impacts through the uptake of work for the targeted breeding system application are summarised in Table 2 and described in more detail below.

Table 2: Prospective impacts from Genomics Phase 2 Digital Initiative - Pan'omics Toolbox (WP1) using CSIRO's triple bottom line (TBL) benefit classification approach

CATEGORY	INDICATOR*	DESCRIPTION
Economic		
Research quality, productivity, and efficiency	↑ Research/ operational efficiency (%) ↑ Quality of research ↓Innovation cycle time/↓Risk (%)	Cost and time savings for the researchers from the uptake of the platform. See (1) in the next section below
New services, products, experiences, and market niches	↑ High-yield (%) and climate-resilient crop varieties (%) ↑ High-yield (%) and climate-resilient livestock breeds (%)	Adoption of WP1 tools to enable the faster lab to market delivery of improved crop and livestock varieties. See (2) in the next section below
National Economic Performance	↑ Revenue from new varieties ($) ↑ income from licenced IP ($)	New income for the adopters from the commercialisation of new products, services and technologies. Potential for export income from new products and technologies. See (3) in the next section below
New Jobs	FTE/ PTE jobs created by the work (#)	The initiative has the potential to create direct, indirect (backward linked industries), and induced employment (forward linkages) requiring a higher skill set in the internal (CSIRO) and external networks
Environmental		
Lower environmental footprint of agricultural activities	↑Crop traits (ability to withstand pests/ diseases etc) ↑Utilisation of available resources (soil/ water/ nutrients) ↑Sustainable crop production	Growing new crop varieties better suited for area-specific requirements (higher yields, disease, and drought resistance) and understanding and controlling pathogens such as invasive species enabled by uptake of new tools to build resistance into otherwise vulnerable species; thereby making them economically and environmentally viable and sustainable for farming.
Conservation of natural plant populations	↓Extinction of crop populations due to environmental stresses (heat/ drought/ frost etc)	Conservation of agriculturally important plant populations that are vulnerable to sudden environmental changes by breeding in adaptability and resilience.
Social		
Health and wellbeing	↑ Societal health	Access to higher quality and healthier food for consumers in adequate quantities.
Food security		
Innovation and human capital (creativity and invention)	↑ Scientific capabilities in this space ↑ SMEs based on new capabilities	Uptake (licenced/bought/open source) of SOI Pangenome Toolbox by biologists and medical researchers to undertake Pan'omics research towards an improved scientific understanding of evolution and to deliver 'omic IP, novel medical diagnostics, and treatments, etc.

*All indicators relate to the adoption of WP1 outputs unless specified otherwise.

Qualitative summary of potential impacts

The uptake of the SOI Pangenome Toolbox and SOI Pangenome Representation outputs will deliver direct benefits from:

1. **Research quality, productivity, and efficiency**

   One of the key economic benefits from the uptake of outputs from Pan'omics Toolbox (WP1) is its potential to improve the research efficiency of breeding programs. It markedly expands the current toolbox available for plant breeding and crop improvement efforts while lowering the overall costs and incorporating time efficiency into these traditionally long and resource-intensive processes. Initial users will be researchers, and end-users will be communities that derive benefits/solutions from understanding the phenotype to genotype relationships, and vice versa, such as breeders or medical researchers.

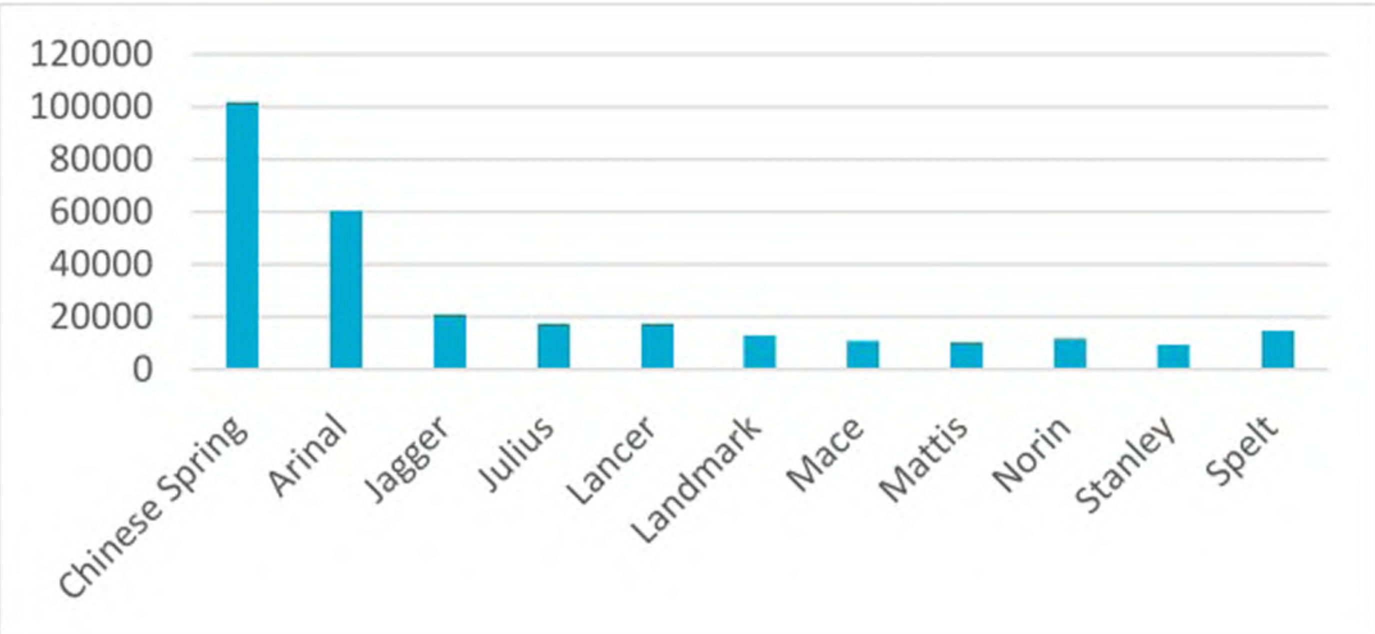
   By acting as an enabler for improving access to the breadth and depth of information/data, the proposed tools have the potential to provide an effective basis for analysis (deeper access to information), fast-track research, create new (research and commercial) opportunities, and address future breeding program research questions and unforeseen problems.

Adopters

Researchers (genomics) in different industries, Agricultural/biotech companies

Quantitative benefits from the research efficiency can be measured as the incremental difference between time and costs expended by the researchers for the two scenarios - "with" and "without" the use of SOI Pangenome Toolbox and SOI Pangenome Representation. Therefore, it also requires considering alternate or substitutable solutions available to the end-users.

Overall benefits will be a function of specific capabilities of the platform, i.e. proprietary knowledge as well as the scale of adoption. It is inherently hard to quantify benefits from new and improved ability to do things that couldn't be done before, or the discovery of new information that wasn't available using the old techniques. The quantitative estimation of impacts would benefit from a survey of researchers/ direct users of the tools; however, this can be expensive to undertake. Some indicators are suggested in Table 2. The impact evaluation approach must be refined once the TRL and Commercial Readiness (CR) of WP1 outputs progress for real-world adoption.



Gain of genomics information as genomes are added to the Annotated-Wheat pangenome, as measured by number of SOI types.

2. New services, products, experiences, and market niches

The uptake of the platform (through licensing, purchase, or open sourcing) by industrial breeding or biotech/medtech companies to undertake Pan'omics driven breeding will be underpinned by access to greater depth and breadth of information. This has potential streamline the breeding pipeline and deliver improved genetically driven gains. The increase in the speed of trait and marker identification will help accelerate the delivery of novel varieties in the form of high-yield and climate-resilient crop and livestock breeds that are better suited to the current global demands. The capability will facilitate the breeding of high-yield, high protein and/or biotic and abiotic stress-resistant crop and livestock species. Similarly, uptake of the SOI pangenome representation by geneticists and tool developers in breeding /biotech /medtech companies will help develop new tools and techniques and advance genomics research.

Adopters

Agricultural/biotech companies, breeders, biologists, and medical researchers

The adoption of WP1 tools has the potential to deliver benefits by aiding the development of new crop and livestock varieties by breeding companies. However, it is important to note that the successful delivery of superior crop and livestock varieties depends upon the combination of several other critical factors. Quantitative benefits from access to new products can only be claimed if the projected impacts (such as new crop and livestock varieties) connect back to WP1 research and innovation activities and the adoption of its outputs. Some indicators are suggested in Table 2 above. The impact evaluation approach must be refined once the TRL and CR of WP1 outputs progress for real-world adoption.

3. National Economic Performance

New sources of revenue such as royalties generated by licencing of intellectual property are expected (scientific discovery at the genome or proteome level, new crop/livestock varieties). Potential revenue and export income from the sale of new crop varieties within Australia and overseas also present key benefit streams as a result of the uptake of R&D.

Quantitative benefits in the form of revenue generated from the sale of new varieties within Australia or overseas. As noted in #2 above, having access to the right tools is only one piece of the puzzle in the development of superior breeding varieties. Hence any claimed benefits must connect back to WP1 research and innovation activities and the adoption of its outputs. The benefits attributable to research are based on consideration of key roles played by collaborating organisations in delivering the overall impact. New income from licensed IP is relatively straightforward to determine. Some indicators have been suggested in Table 2. The impact evaluation approach needs to be refined once the TRL and CR of WP1 output progress for real-world adoption.

4. Improved capability

The breeding programs would train the next generation of researchers and plant breeders in optimal utilization of WP1 outputs. This has the capability to catalyse innovation and provide a great source of value to Australia. The initiative has the potential to add to the general body of knowledge, which may contribute to other discoveries and development while creating options across a wide range of research areas, new SMEs and industries. Some of the potential benefits from these new options emerge from augmenting national science and technology capability, improved knowledge, better research infrastructure, improved industry benchmarks, a clearer understanding of the most fruitful areas for future research, and information on what areas of research might best be prioritised/ scaled-back/abandoned. Many of these are not readily evident at this stage. Considering the depth and breadth of CSIRO's work, and the gamut of generated options, it is impracticable to assess these benefits within the current scope of this impact assessment.

Key assumptions

The adoption costs associated with sequencing and assembling a pangenome is a key factor that will determine its impact potential. WGS, which forms the bedrock of an SOI Pangenome, is undergoing an exponential decrease in cost but is not commercially viable for population-scale genomic analysis, thereby limiting the extent to which it can be used. For example, performing WGS on every individual in a breeding program is cost-prohibitive but tractable for the founders. Cost is also related to size and so efforts to sequence simpler pathogens (e.g. wheat rust) at larger scale are also tractable.

DNA sequencing has evolved over three decades of incremental technological advancements.[14] In parallel, there has been an exponential decrease in the associated cost of DNA sequencing. For perspective, the sequencing costs have fallen from $0.52 per Mb of DNA sequence in 2010 to just $0.010 in 2019. This combination of constantly improving technology and decrease in associated cost has acted as an enabler for widespread adoption of the technique.

The impact assessment is based on the premise that WGS costs will continue to follow a trajectory of exponential decline which is likely to drive up the adoption rate. As the cost of WGS decreases and population-scale genetic analysis through pangenome creation becomes common, the cost-efficiency of WP1 incorporation as an essential breeding toolbox will increase proportionately. However, if the technology adoption costs remain high, the application will remain limited to niche markets and high-value breeding programs alone.

Key impact inhibitors

- **High full genome sequencing costs:** Adoption rate impacted by slower than anticipated decline in WGS costs. See above
- **Failure of R&D hypothesis for use of Toolbox for functional analysis and/or genomic-driven breeding:** Although the scientific rationale behind the Pangenome construction and development of WP1 indicates that the collection of structural information should lead to greater insights, this is yet to be validated through further research. There remains, therefore, potential risk of the limitation of the SOI Pangenome and WP1 as a useful technique for functional analysis and genomic-driven breeding. Nonetheless, it would still remain relevant to and valuable for comparative genome analysis-based applications.
- **Competitor products:** WP1 R&D will remain under development for at least the next 2-3 years before starting to work on a commercialisable tool. Based on inputs from the team, there are no competitor products at this stage. However, considering the low maturity of technology, R&D delays associated with lack of a continuous investment stream, and an expected longer-term period for commercial tools to deliver real world benefits, there is an inherent risk of another organisation (domestic or global) delivering a similar or better disruptive tool. This could jeopardise or delay the delivery of impact.

[14] The sequence of sequencers: The history of sequencing DNA - PMC (nih.gov)

7 Video Phenomics (WP2) - Qualitative summary of prospective impacts

Box 4 – Video Phenomics (WP2) rationale

Industry Gap

Crop yield is a function of dynamic interactions between genome, environment, and management. However, the current crop breeding programs are limited by the dearth of dynamic information on interactions to influence the performance of a particular genotype. The collection of consistent and accurate phenotypic data is challenging, labour and time intensive, sometimes destructive and requires a lot of experience. Farmers currently rely on demanding and destructive methods to evaluate crop performance. Breeding teams have reported spending as much as 50% of their time in collecting this information and the quality of data still remains prone to human error.

Hence the industry has a need for reliable and non-destructive ways to measure and monitor crop performance in the field of phenotypes over time to better inform selection decisions.

Prospective incremental change

WP2's artificial intelligence-based Video Phenomics platform has potential to offer a non-destructive, accurate and high-throughput digital solution for crop performance monitoring and yield estimation to improve the efficiency of crop assessment at the farm scale. The adoption of such a tool has the potential to provide greater volume, consistent, and accurate data to allow easier identification of traits for improvements in crop varieties and allow uniform evaluations that are less susceptible to human bias. This ML powered platform helps with automatic capture, quantification, and prediction of traits to enable accurate collection of phenomics data, such as emergence, spike and spikelet counts, height and biomass index, and to correlate this information with growth performance parameters such as yield. This results in additional operational efficiency and contributes towards improved farmer decision making.

Although solutions to tackle one or more subsets of this problem have emerged in recent years to varying extents both within and outside Australia, there currently are no known examples of end-to-end integration and digitation of phenotyping. Provision of an end-to-end platform through this effort is a key scientific differentiator compared with other solutions available to end-users. The work also has the potential to play a significant role in broadening the accessibility of phenomics technology and increasing its scalability.

Key Trends

Highly accurate, automatic, multifunctional, and scalable phenotypic technologies are widely recognised globally as important tools for accelerated advancement of genetic gain in breeding programs. Video phenomics has the potential to provide breeding scientists with new insights to provide intelligent solutions for breeding programs.

To exploit genomic breeding, there is a need to evaluate environmental growing conditions, by using advanced imaging, drones and satellites, as well as new tools to measure desirable (and undesirable) plant qualities. Recognising the limitations of phenotyping efficiency, the 21st century has seen considerable research efforts globally focussed on development of improved phenotyping solutions.

There is a radical demand for commercial and research tool solutions that are capable of collecting massive amounts of phenotypic data from hundreds of plants every day with a high degree of automation.

Prospective Impacts

The potential TBL impacts through the uptake of work are covered in Table 3 below.

Table 3: Prospective impacts from Genomics Phase 2 Digital Initiative - Video Phenomics (WP2) using CSIRO's triple bottom line (TBL) benefit classification approach

CATEGORY	INDICATOR*	SUPPORTING OUTCOMES
Economic		
Efficiency and Productivity	↑Operational efficiency(%) ↑ Safe work at farm ↓ Labour costs ($) and ↑time savings(%) ↑Field time utilisation (%)	Improved operational efficiency and data reliability from the uptake of video phenomics platform; a shift from manual to automatic phenotypic data collection. Accelerated on-farm collection of real-time and in situ data, curation, and real-time visualisation of trait data at scale from the deployment of AR tools at the farms. See (1) in the next section below
Research quality, productivity, and efficiency	↑ Quality of research (phenotyping)	
Management of risk and uncertainty	↑Accuracy and consistency of collected trait data (%) ↓ Risks and uncertainty (%) ↑Improved management (%)	Improved breeding program management and high confidence decision-making enabled by access to new tools for farm management (AR powered) and real-time data that is less labour intensive. More timely interventions leading to reduced costs/ impacts. See (2) in the next section below
New services, products, experiences, and market niches	↑ High-yield (%) and climate-resilient crop varieties (%)	Uptake of data furnished by WP2 tools for research by breeders to enable the delivery of improved crop and livestock varieties. See (3) in the next section below
National Economic Performance	↑ Revenue from new varieties ($) ↑Crop yields (%) ↑ Income from licenced IP ($)	New income for adopters from the commercialisation of new products, services and technologies as discussed above. Potential for export income from new products and technologies. See (4) in the next section below
Environmental		
Lower environmental footprint of agricultural activities	↑Crop traits (ability to withstand pests/ diseases etc) ↑Utilisation of available resources (soil/ water/ nutrients) ↑ Sustainable crop production	Growing new crop varieties better suited for area-specific requirements (higher yields, disease, and drought resistance) enabled by uptake of new tools to build resistance into otherwise vulnerable species; reduced use of insecticides, thus making them economically and environmentally viable and sustainable for farming.
Conservation of natural plant populations	↓Extinction of crop populations due to environmental stresses (heat/ drought/ frost etc.)	Utilising more timely information about species and better evaluating threats to them through climate change, pathogens, etc. Conservation of agriculturally important plant populations that are vulnerable to sudden environmental changes by breeding in adaptability and resilience.

CATEGORY	INDICATOR*	SUPPORTING OUTCOMES
	Social	
Health and wellbeing	↑ Societal health ↓On-field manual data collection	Adoption of WP2 has the potential to allow a better lifestyle for agronomists and field scientists through fewer trips to the field for data collection and better field time utilisation.
Food Security	↑ Crop yield ↑ Food production and quality	Access to higher quality and healthier food for consumers in adequate quantities. Feed the growing population by growing food sustainably
Innovation and human capital (creativity and invention)	↑ New frontiers in breeding strategies	Uptake of the Video Phenomics platform provides a base to advance phenomics technology to better support sustainable agriculture and spur the growth of SMEs in this space while leveraging existing and new partnerships.

*All indicators relate to the adoption of WP2 outputs unless specified otherwise.

Qualitative summary of potential impacts

1. Improving operational efficiency and productivity

The access to non-invasive, high throughput and accurate data delivered by the ML-powered video phenomics platform has the potential to significantly improve the upstream and downstream operational efficiency of breeding processes through alleviating longstanding bottlenecks (data quality and time) associated with traditional manual operations. It improves data collection as well as increases the capacity to extract meaningful information from available data to better support downstream operations thereby providing a potential to decrease cycle time to improve crop varieties. It also provides better and more informed remote connectivity between the farm and farmers/breeders, thus facilitating improved remote farm management.

Automation of data collection will reduce the labour costs while empowering practitioners to channel time and money into focusing on solving more pressing problems rather than expending these resources on obtaining the field data. Understandably the adoption of automated solutions also has the potential to contribute to safer farm practices through lowering the need for manual operations and exposure to farm chemicals.

Quantitative benefits from the operational efficiency can be measured as the incremental difference between time and costs expended by the researchers for the two scenarios – “with” and “without” the use of the Video Phenomics platform. Hence it also requires considering to alternate or substitutable solutions available to end-users to address the data collection requirements.

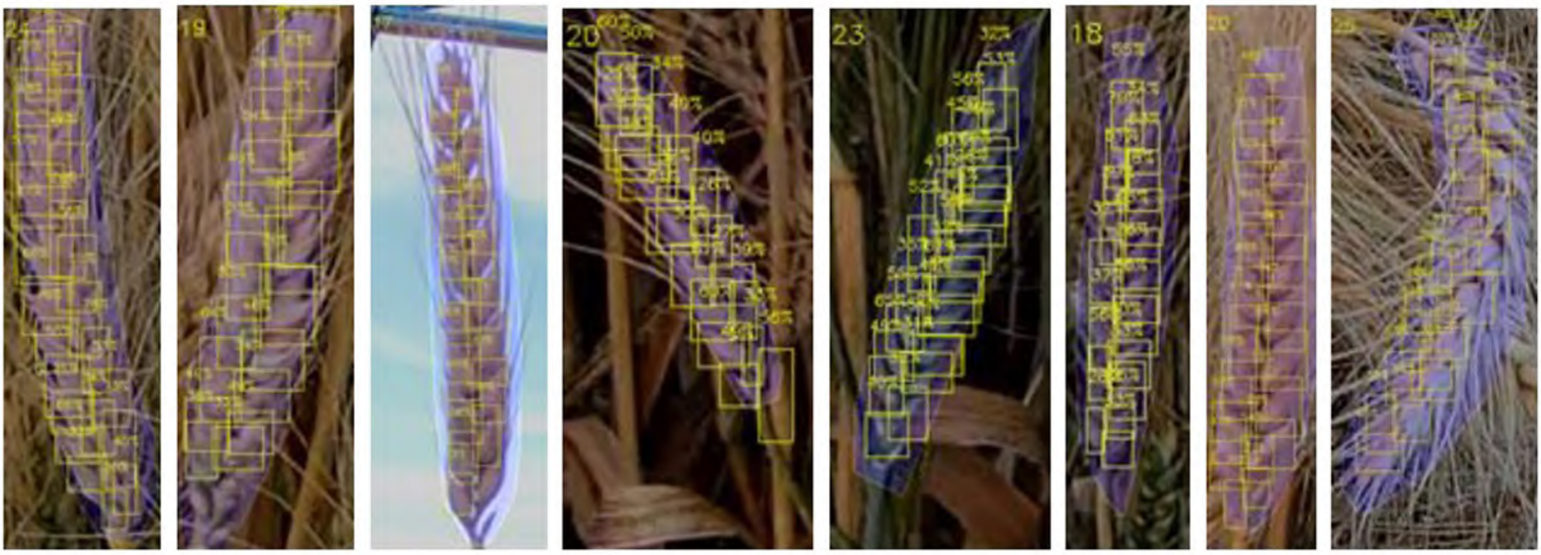
Overall benefits will be a function of critical capabilities of the platform (such as accuracy, overall throughput, analytics, the practicality of use, scalability etc) as well as the scale of adoption. The quantification of impacts would benefit from a survey of researchers/ direct users of the platform; however, this can be expensive to undertake. Some indicators are suggested in Table 3. The quantitative benefits from operational efficiency that contributes towards shorter time to market or safer farm work practices are difficult to estimate at this stage. The impact evaluation approach must be refined once the TRL and CR of WP2 outputs progress for real-world adoption.

2. Management of risk and uncertainty

The deployment of (licensed/bought) of the video phenomics platform/ AR tools in the field by industrial breeding (e.g. InterGrain) or agriculture survey companies/ farmers will facilitate conducting and accelerating on-farm collection, curation, and real-time visualisation of trait data. The new farm management tools, greater volume, speed, quality, and consistency of data combined with lower susceptibility to human bias/errors, may assist with uniform evaluations to support improved and high confidence decision-making and crop performance predictability.

The adoption of WP2 tools has the potential to provide better data that can feed into improved management of risk and uncertainty. However, it is important to note that any benefits from improved management are dependent upon a combination of several other critical factors; good data only forms a part of the overall requirements. Hence any quantitative benefits from improved management can only be claimed if the changes connect back to the uptake of WP2 outputs. Some indicators are suggested in Table 3 above. The impact evaluation approach to assess benefits from the management of risk and uncertainty must be refined once the TRL and CR of WP2 output progress for real-world adoption.

3. New services, products, experiences, and market niches



Sample spike images with detected spikelets. yellow bounding boxes represent the detected spikelets and the number at the top left corner of each image represents the spikelets count

The uptake of WP2 tools by breeders for collecting and analysing phenotypes and understanding complex relationships between trait expression and environment has potential to deliver superior performing breeding varieties (high-yield/climate-resilient) more suitable for a particular soil, weather type, etc.

The adoption of WP2 platform has the potential to deliver benefits by aiding the development of new crop and livestock varieties by breeding companies. However, as mentioned above in WP1, the successful delivery of superior crop and livestock varieties depends upon the combination of several other critical factors. Quantitative benefits from access to new products can only be claimed if the projected impacts (such as new crop and livestock varieties) connect back to WP2 research and innovation activities and the adoption of its outputs. Some indicators are suggested in Table 3 above. The impact evaluation approach must be refined once the TRL and CR of WP2 outputs progress for real-world adoption.

4. **National Economic Performance**

   Same as for WP1, see above. For indicators see Table 3.

5. **Improved capability**

   Same as for WP1, see above. For indicators see Table 3.

   In addition, the platform also has other compelling use-cases in addition to the breeding program. For example, the platform can potentially be deployed to detect adverse conditions on a farm, such as drought stress, signs of pathogen effects, etc. This and other similar use cases have the potential to deliver TBL impacts but not readily evident at this early stage of technology maturity.

People at a research farm with AR devices including HoloLens glasses, phone and iPad (left); and an example of plot specific digital holograms floating above the respective plots (right)

Key assumptions

The projected benefits are based on a significant presumption that the WP2 Video Phenomics platform is capable of providing accurate and reliable data and does not suffer biases similar to manual methods. The testing of the platform recently has been carried out by CSIRO Data61 and Ag&F research groups using the data collected at a CSIRO agricultural research station as well as an industry partner - InterGrain's breeding sites under the scope of SIEF Genomics P2. Although CSIRO trials delivered acceptable accuracy, the testing of the platform at InterGrain sites highlighted key issues with blanket adoption and applicability of the work package. This underlined the need for further development and testing of the platform as well as end-user training requirements to be able to utilise the platform for targeted applications. It is evident that the impact potential of the platform is directly dependent on its effectiveness, accuracy, and usefulness. Some of the impact risks are discussed below.

Key impact inhibitors

- **Lack of technical prerequisites at deployment sites:** Optimal performance of WP2 requires adherence to technical parameters (e.g. speed/ height of the UAV/ angle of the camera/ frame rate of video capture etc.) to deliver acceptable data quality as well as the accuracy of predicted traits. For example, when capturing videos through cameras mounted on moving vehicles/UAVs, high frame rate can help reduce the blurriness of the video. There is also an inverse relationship between the video frame rate and image resolution. Some or all of these prerequisites are not always optimally met at deployment sites.
- **End-user site limitations:** Successful commercial deployment of WP2 and the ability of software to deliver sufficient accuracy is contingent upon customer-specific field conditions. For example, it is possible that the end-users (eg. breeders) may be bound by certain topographical limitations that make adherence to prescribed hardware configuration parameters challenging. These kinds of factors can impact the quality of image/ video data collected, and therefore limit the predictability function of the platform.
- **Adoption costs:** High end to end usage costs could be a key inhibitor in the pathway to adoption of the technology.
- **Lack of applicability across different crop systems:** Large-scale adoption of WP2 may be limited by its lack of application across multiple crop systems. This is an inherent risk in all ML-based categorisation tools. It requires software tailoring to be able to recognize traits relevant to each species with reasonable accuracy. For e.g. the current testing and development of WP2 has been based on wheat as the model species; the transferability of the software has been tested on Barley and oats. While transferability potential to Barley delivered satisfactory results, the platform performed sub-optimally on oats where the shape and morphology of spikes is quite different from wheat.

  In addition to this, irrespective of the model system, the video phenomics platform detects traits using the image and video data collected from fields, there are always some traits that may be missed due to occlusion. However, a correlation between the number of detected traits and the actual number of traits can be established by using adequate mechanisms such as regression modelling. This needs to be developed for each application.
- **Inability to perform timely validation and optimisation of WP2 outputs:** For breeding programs, phenotypic data collection is time-bound by specific phases in a plant's growth cycle. Missed opportunity requires researchers/ breeders to wait for the next growth cycle. This is applicable while developing training data for any given species. Although this risk is inherent to the domain, instead of being a limitation in the technology it does underline the need to capture data for at least two crop life cycles to obtain statistically significant results. The quality of the training set is a determining factor in the accuracy of predictability of a developed model. However, once the model has been trained on a reliable training data set (replicated life cycle), it should be applicable to any generation.

  If the lack of funding and support results in a missed crop cycle, it could delay the software optimization for another crop cycle as well as its subsequent-technology deployment whilst extending overall project costs and timeline (including costs associated with instrument maintenance while waiting for the next cycle).

8 Recommendations and support requirements

A combination of top-down and bottom-up approaches is required to channel the current outputs from WP1 and WP2 towards potential future impact through development from a product perspective and identifying future markets/use cases. Some recommendations based on discussions with the core team include:

Recommendations

- **Market Analysis:** Engagement with market analysis and associated product development teams would assist in charting a realistic and practicable pathway to commercialisation and the adoption of potential future outputs from these work packages. A local and global market research, as well as customer surveys with the support of CSIRO Business Development and Commercialization teams, could potentially provide valuable market insights such as customer interest, size of the potential market, key competitors, roadblocks, the current standing of the tools in the marketplace, their potential high-value applications and relevant collaborations to be pursued. The effort can also play an instrumental role in identifying the most feasible pathway to impact.
- **Focus on shorter-term outcomes while targeting longer-term impacts:** Overall, the WP1 is at a relatively lower TRL level (concept stage). The comparative genomic analysis functionality, however, has been validated and upon project funding continuation, the team envisages the availability of a tool that forms a core capability to produce, analyse, and visualise SOI-based pan genomes in the next 3 years. The operational genomics functionality is likely to incubate for another 4-5 years before delivering any commercial benefits on investment. During this period, increased visibility through publications and presentations will be one of the avenues to garner potential collaborator interest. Any real-world impacts are not expected to be realised before the next 5-10 years atleast. To continually progress on the TRL and CR readiness it is important to focus on shorter-term advancements while keeping the longer-term impact goals in sight.

  Similarly, for WP2 collaborations with the crop industry including crop breeders and crop survey companies, and identifying early adopters for uptake of the video phenomics platform can help test and validate platform features for continuous improvement as well as assess its usability, readiness and customer interest. To summarise, keeping the needs of key customers/ early adopters such as research communities and breeding companies at the heart of technical interventions, as well as their early engagement by way of a robust uptake plan from the outset, is critical to realise intended impacts.
- **Stakeholder engagement and communication:** Engaging with the potential end-users through channels such as industry workshops, industry and marketing events, conferences etc, can help build meaningful networks for the commercialisation of the platform. To obtain industry traction, the value proposition of the work needs to be clear and compelling when compared with current practices. This would help build confidence and propel adoption to deliver and scale impacts.
- **Impact thinking and planning from the outset:** It is recommended that the focus on commercial adoption and building compelling use cases with the uptake of the WP outputs is as strong as delivering scientific impact. The lack of focus on impact planning and tracking (around how the new interventions are being advanced and/ or adopted to deliver impact) from the outset presents a key gap in the innovation system, due to which research outputs are frequently unable to effectively realise their impact potential.

Support Requirements

Some areas of immediate support requirements from CSIRO to progress on the impact pathway, based on the inputs from the R&D team include:

- **Leadership support:** Executive sponsorship is required to drive scientific development as well as the proposed bridging activities within CSIRO for uptake of WP outputs for potential applications in other research projects, as discussed in Outcomes section above. The lack of direct alignment of the initiative with CSIRO missions might act as an impediment and discourage the decision-making body from investing in further development of these technologies. It will be useful if CSIRO management engages with the program, to evaluate its potential and make decisions on funding further development.

It is also recommended that the team provides a thought-through, clear and coherent articulation of support requirements to the leadership team for this large-scoped and longer-term initiative. An understanding of the expectations early-on will help add efficiency to the innovation cycle; translate R&D into best practice solutions and increase the probability of all stakeholders benefitting from enhanced capability.

- **CSIRO internal collaborations:** As noted above, involvement in CSIRO projects that offer bridging activities to tie the needs and requirements of these projects with the capabilities of WP1 and WP2 to enhance their research outcomes have potential to support the testing, validation, and advancement of tools. Collaboration with groups that collect genomic information and/or CSIRO mission programs provide attractive avenues for mutual benefit for the groups involved.

- **Funding:** With the completion of SIEF funding, both the R&D teams require new sources of dedicated funds and resources to pursue the TRL and CR advancement of the respective WP outputs.

  - **WP1-** Where feasible, short-term investments can yield returns such as articulating the significance of the tool and its value-creation to the research community through enhancing their understanding of the genomic data at hand.

  - **WP2-** Continued communication with the collaborators can help optimize the software workflow and improve the user interface for the end-users. InterGrain has shared their data for assessment and analysis. They have shown continued interest in the adoption and deployment of WP2 and provide feedback for further improvement of the video phenomics platform. This has potential to improve WP2 software capability expansion as well as optimisation and eventual deployment of the video phenomics platform for Wheat.

  However, any future efforts for advancing WP (1&2) outputs necessitate the acquisition of dedicated funding sources.

- **Resources:** There is a need for cross-disciplinary skills and dedicated resources to advance the TRL and CR of both work packages. Some of the critical support requirements include:

  - Internal scientific support from the Data 61 team for statistical analysis and modelling towards developing capabilities of WP1 for functional genomics applications.

  - The expertise of product engineering units to expand the comparative capability of WP1 to similar datasets such as transcriptomes and proteomes.

  - Upscaling the WP2 platform for large-scale survey and expanding the functions of the platform to include more use cases requires dedicated software engineering resources.

  - Support from CSIRO business units (BUs) especially Agriculture & Food to determine the best commercially relevant use cases for the tools while integrating the expertise of the product and design group to drive product development accordingly.

Evaluation challenges

The quantification of benefits is generally conducted using CBA. The method compares the projected benefits of a project against its costs to provide a Benefit Cost Ratio (**BCR**) and Net Present Value (NPV).[16]

The costs considered include the costs incurred by CSIRO and its research partners to produce the research outputs in the chosen assessment period. Where data is available, usage and adoption costs borne by end-users should be included. The benefits calculated in the analysis are the net benefits from the program, that is, the difference between the 'with program' and 'without program' scenarios.

As evident from the discussion noted above earlier in the assessment, with WP1 at TRL 3 and WP2 at TRL 3-4 – there is uncertainty with respect to both – the overall costs (developmental and adoption) as well as the extent of benefits with the adoption of Outputs. Many aspects of both the work packages require further testing and validation. Hence a quantitative assessment is not practicable at this stage.

[16]The formula for calculating a benefit cost ratio is defined as:

Benefit Cost Ratio=PV(B_t)/ PV(C_t)
Net Present Value=PV(B_t)-PV(C_t)
Where
PV(B_t) is the present value of the benefits at time t
PV(C_t) is the present value of the costs at time t

9 SIEF'S role

SIEF funded the Genomics Phase 2 Digital Initiative through the Future National ICT Industry Platform Program enabled by the NICTA Gift, to further develop the outputs delivered in the previously funded Genomics Phase 1 Digital Initiative. The purpose of the FNICTIPP is to support research activities in the field of information and communications technology (ICT) that benefit Australia by helping create new Australian technology-based industries and/or applied technology platforms that can reach a global scale.

SIEF has played an instrumental role in laying the scientific groundwork to develop technologies and software platforms under the Genomics P2 initiative to deliver prospective impacts covered above in this report. The support has provided a significant basis for further scientific and financial interventions as well as collaboration decisions for the future development of the WPs. The initiative provided a platform for research integration of agronomy, life sciences, information science, math and engineering sciences, and it combines high-performance computing and artificial intelligence technology to develop new methods of mining genes associated with important agronomic traits and propose new intelligent solutions for breeding programs.

Future evaluations post the TRL, and CR advancement of R&D must acknowledge attribution of the SIEF investment outputs originating from this initiative as inputs into their research pathways to impact.

10 Conclusions and confidence rating in the impact assessment

The Genomics Phase 2 Digital Initiative initiative is aspirational. Although the adoption of WP outputs into the breeding programs are the key focus at this stage, the work has the potential of transferability and value creation across different sectors. The initiative provides a good example of R&D projects that may not produce immediate commercial results but have the potential to push scientific frontiers to drive strategic crop research to deliver against the challenge of Food Security.

This ex-ante assessment is an early-stage analysis based primarily on the discussions with the project teams. The successful delivery of impacts is highly contingent upon an interconnected mesh of technical, market, regulatory, and economic factors, many of which are unpredictable at this stage, due to the nascent nature of the industry.

At the time of writing this report, the R&D work for both WPs has paused with the ending of the SIEF support program in May 2022. In order to progress on the impact pathway, the next phase of funding support is critical. It is also important to have clarity on the alignment of this work with CSIRO's strategic priorities while keeping industry interest and global developments and trends in mind.

The projection of benefits is based on several hypotheses and hence the confidence rating in this impact assessment is rated very low by Tractuum. If and as CSIRO's research progresses, the current study should be revisited and refined, drawing on more detailed evidence (including the external stakeholders e.g. InterGrain) to provide greater insights to inform decision making.

Appendix A Significance of impact management

Globally there has been an unprecedented focus on demonstrating impacts generated from R&D investments; however, little effort has been put into establishing baseline information by which it can be measured. **A recent survey suggests that early impact measurement and management practices are key contributors to market growth; 88% of respondents reported that it drove better than expected returns on financial investments, and 99% reported meeting or exceeding their impact expectations.**[19] This provides an additional boost to the incorporation of impact estimation, planning, monitoring and evaluation frameworks as a part of project activities.

Some leading organisations globally track and report their impact across the economic, environmental, and social landscape.[20, 21, 22] This demonstrates to key stakeholders the positive and negative 'shifts' which the activities of the consortium have delivered. Well-executed impact management and reporting help create brand affinity, accountability and transparency, industry leadership, and a solid strategy roadmap.

Impact assessments require structured and coordinated measurement for the benefit of key stakeholders to estimate and monitor progress (investors); benchmark investment effectiveness (portfolio managers); measure progress (enterprises or investees); and catalyse adoption (beneficiaries). **Impact measurement efforts provide a useful resource to prioritise and plan research expenditure.**

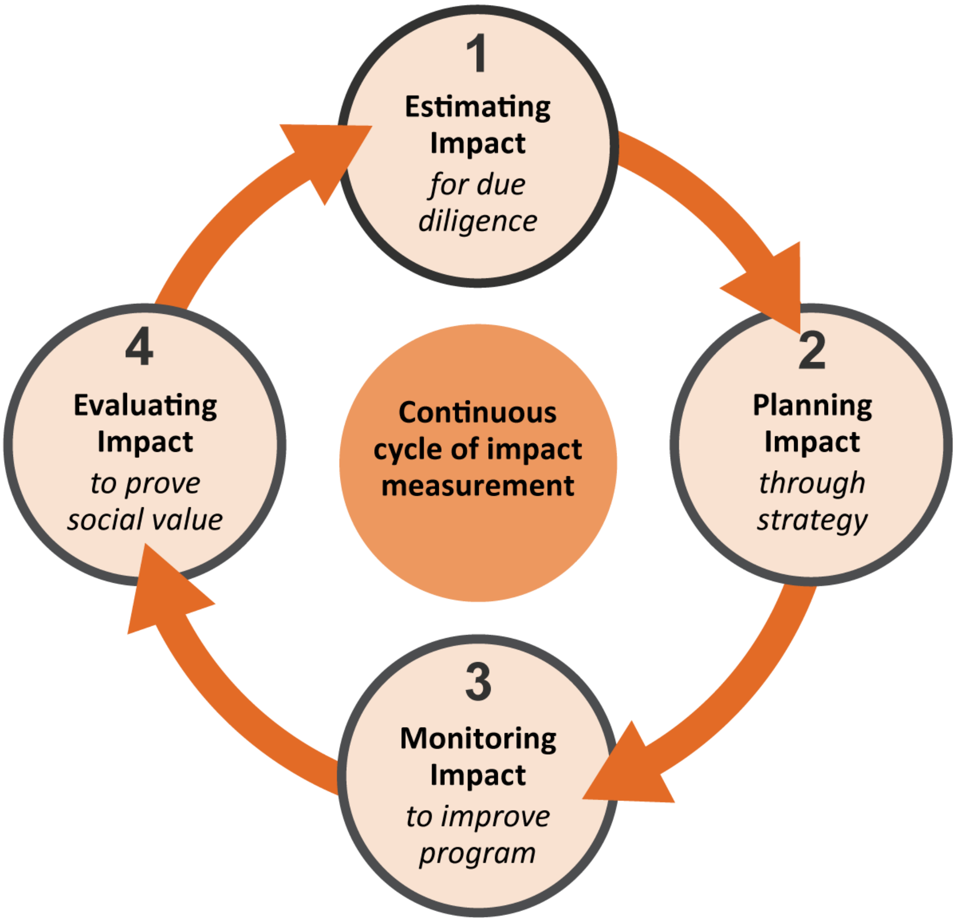


Figure 5: Continuous cycle of Impact measurement objectives[23]

[19] https://www.pionline.com/esg/global-impact-investment-market-going-strong-despite-pandemic-giin
[20] S&P Global Impact Report
[21] Shopify Impact Report
[22] Diabetes UK Impact Report
[23] www.hbs.edu/socialenterprise/Documents/MeasuringImpact.pdf

Appendix B Employment contributions

It should be noted that any additional employment (typically stated as 'jobs created') is not an economic benefit. Just as for any other resource, the use of additional labour resources imposes an opportunity cost on Australian society, because those workers cannot be used elsewhere to produce goods or services. In addition, some workers will simply transfer from other jobs (potentially including CSIRO positions), so the net creation of jobs will be zero. Those workers who are employed in new positions will obtain a wage, but the cost of the wage is borne by employers, so the net benefit to society is zero, except for any additional profit (producer surplus) that is generated. **Nevertheless, estimates of job creation opportunities are generally of interest to decision-makers, and they can be reported separately from the cost-benefit analysis to provide a comprehensive outline of expected impacts.**

In principle, the engagement of an unemployed worker with no other clear job prospects imposes no opportunity cost on society. **In a situation of structural (i.e., non-cyclical) unemployment, therefore, society can benefit from the creation of new jobs that are filled by the unemployed. But this benefit can only be realised if the skills of the currently unemployed workers match the competencies required in the newly-created jobs.** Further, any benefit to the newly employed workers, and hence society, would be offset to some extent by their loss of leisure (i.e., non-work) time, which can also result in social benefits through activities such as child-minding, gardening, relaxation, exercise, etc., that are valued by the worker.

Taxes have a depressive effect on the economy by reducing aggregate demand and/or output. They, therefore, reduce job opportunities and profits. To the extent that the SIEF Genomics Phase 2 Digital Initiative is funded by CSIRO, SIEF and other funding sources through government taxation, there will be some potential loss of jobs in the economy. In other words, it cannot be claimed without qualification that there will be a straightforward increase in employment levels attributable to the assessed work.

Appendix C References

1. www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/HLEF2050_Global_Agriculture.pdf
2. https://www.mdpi.com/2073-4395/8/12/277/htm
3. https://www.frontiersin.org/articles/10.3389/fpls.2019.00714/full
4. https://www.health.gov.au/initiatives-and-programs/genomics-health-futures-mission
5. https://www.theworldcounts.com/challenges/planet-earth/state-of-the-planet/is-the-world-running-out-of-food/story
6. https://www.embopress.org/doi/full/10.1002/emmm.201100153
7. https://www.frontiersin.org/articles/10.3389/fmicb.2016.01488/full
8. https://sief.org.au/nicta-gift/future-national-ict-industry-platform-program/genomics-digital-initiative/
9. www.hbs.edu/socialenterprise/Documents/MeasuringImpact.pdf
10. The sequence of sequencers: The history of sequencing DNA - PMC (nih.gov)
11. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7838587/
12. https://www.pionline.com/esg/global-impact-investment-market-going-strong-despite-pandemic-giin
13. S&P Global Impact Report
14. Shopify Impact Report
15. Diabetes UK Impact Report