Final Report

National pineapple breeding and evaluation program

Project leader:

G. Sanewski

Report authors:

G. Sanewski

M. Webb

N. Pathania

Delivery partner:

Department of Agriculture and Fisheries

Project code:

PI17000

Project:

National pineapple breeding and evaluation program (PI17000)

Disclaimer:

Funding statement:

This project has been funded by Hort Innovation, using the pineapple research and development levy and contributions from the Australian Government. Hort Innovation is the grower-owned, not-for-profit research and development corporation for Australian horticulture.

Publishing details:

Published and distributed by: Hort Innovation

Level 7
141 Walker Street
North Sydney NSW 2060

Telephone: (02) 8295 2300

www.horticulture.com.au

Contents

Public summary

The Department of Agriculture and Fisheries (DAF) has now completed two phases of breeding of fresh market pineapple. The first phase focused on selecting varieties with improved eating quality and resistance to translucency in comparison with the major fresh market variety, Smooth Cayenne. The second phase aimed to improve yield without compromising on eating quality. The intent was to develop pineapple varieties suited to the main summer production season in tropical and subtropical production regions of Australia. Over the course of the 5-year National Pineapple Breeding and Evaluation Program, over 35,000 seedlings were fruited across Queensland pineapple growing districts including SE Qld, Wide Bay, Coastal Central and Far Nth Qld. A total of 335 selections were made from a broad range of parent combinations. Planting material was retained from all selections and a low level of multiplication achieved in most cases. Selection was biased towards strong, vigorous plants with good fruit size and good eating quality. It is expected there will be several selections with good potential suited to each of the regions.

Further work on molecular markers was also completed to fast track the final selection process of the preliminary selections. A replicated field trial examining resistance to natural flowering over the winter of 2022 was completed. Genetic analysis identified an associated position on chromosome 4 and a causal gene proposed. Re-analysis of previously collected data on resistance to natural flowering from drought identified an associated position on chromosome 6 and a causal gene proposed. This work now allows the use of molecular markers to assist the final selection process.

31 varieties held in the DAF pineapple germplasm collection were re-sequenced using the Illumina platform to produce a high quality, high density set of SNP markers to assist with improved genetic analysis. DArTseq markers sets held for approximately 1,000 accessions, were also re-referenced against several new pineapple genomes to allow more accurate positioning of genes of interest. This work improves the accuracy and usefulness of our pineapple genetic analysis tool kit.

Plant and fruit data was collected on a diverse population of genotyped seedlings and varieties to identify molecular markers associated with many traits of importance other than just resistance to root disease and natural flowering, including fruit size, sugar and acid content, flesh color, core diameter and slip number. This data now allows marker-assisted selection and genomic selection for many traits of interest to be included as breeding tools in pineapple.

DAF now holds a very good germplasm collection complete with high quality genomic data to allow further gene identification and improved breeding outcomes. In addition, a large population of new selections are now held ready for further multiplication, selection and testing in growing regions.

Keywords

Pineapple, Breeding, Varieties, Molecular markers, GWAS, Genomic selection

Introduction

The Australian pineapple processing sector has seen a significant decline over several years, principally due to competition from cheaper imported product. The high cost of production and processing in Australia has meant reduced profitability for the principal processor, Golden Circle - Heinz Ltd. Changes in ownership of Golden Circle - Heinz Ltd now allow a greater component of imported product to be used. The decline in production for the processing market has been compensated for, to some degree, in recent years by an increase in production for the fresh market. The Australian pineapple industry is now predominately a fresh market industry with a value equivalent to 92% of the overall pineapple production.

The fresh sector was protected from competition up to 2004 due to a quarantine ban on the importation of fresh product. This sector has expanded considerably in the last 12 years, principally due to the adoption of better fresh market varieties. However, several countries including Thailand, Malaysia and the Philippines can now import fresh pineapple into Australia. It is difficult to predict the future quantities of fresh imports or their likely impact on the Australian industry.

Into the future, the fresh sector should be better placed to compete with imported product compared to the processing sector, principally due to product freshness and consumer loyalties but nevertheless it needs to ensure its position and improve profitability. Profitability in fresh pineapple production has declined slightly in recent years as costs of production rise. Major improvements in product quality, continuity of supply and efficiency of production are necessary.

Pineapple is unique amongst fruit crops in that it is produced all year. In Queensland, it is grown from southern Queensland to Far North Queensland. In an ideal world, the program would deliver a single variety with suitability across a range of climatic conditions. However, this is not realistic due to the wide variation in growing conditions and disease pressure across regions. Therefore, the program focused on developing a range of varieties with improved production characteristics and suitability to different environments.

The Department of Agriculture and Fisheries (DAF) has now completed two phases of breeding of fresh market pineapple. The first phase focused on selecting varieties with improved eating quality and resistance to translucency in comparison with the major fresh market variety at the time, Smooth Cayenne. The second phase aimed to improve yield without compromising on eating quality. Opportunity exists to further improve on these varieties by incorporating greater resistance to Phytophthora root rot (Pc) and Natural flower initiation (NI).

The aim was to develop pineapple varieties suited to the main summer production season in tropical and subtropical production regions of Australia. The intent was to develop varieties superior to the current main commercial varieties 73-50 and MD-2 in yield, disease resistance, plant vigor, crop control and eating quality. For the first time, the traits of resistance to *Phytophthora cinnamomi* (Pc) and NI were incorporated, and molecular markers were developed to assist selection. An emphasis was placed on using parents likely to impart resistance to Pc and NI. Several varieties in the DAF collection appear to have a moderate to good level of resistance to either Pc or NI. Pc resistance appears to be a dominant trait and incorporating a moderate to

good level of resistance into new varieties while retaining other quality attributes was considered achievable.

It was expected that incorporation of Pc and NI resistance would improve ratoon cropping and drought tolerance. Similarly, incorporation of low slipping potential was expected to reduce the cost of slip removal. The program also sought to obtain yield potential better under all environmental conditions than the current industry standards. Other listed traits were expected to reduce post-harvest rejects and improve eating quality.

However, domestic pineapple varieties are highly heterozygous, and it is considered difficult to improve multiple traits simultaneously. When using traditional breeding approaches, efficiency is typically low, and several generations of large populations are needed to combine several traits of interest into a desirable elite single genotype. The challenge is to reduce the interval between generations, reduce population sizes and improve the concentration of desirable genes in parental combinations.

The inter-generational interval is usually 7-13 years in pineapple because of the need to multiply and test varieties in G X E trials and to measure their breeding value before they can be used as a valuable parent for a subsequent generation. The program considered ways to reduce this interval to effectively increase the number of generations and present new desirable varieties to consumers in a shorter time.

Progeny populations in pineapple are generally in the order of 20,000 plants or more. It is expensive to develop and manage a population of this size. The program sought out an approach to reduce the necessary population size to a manageable but effective level, minimize operational costs, and minimize the reliance on phenotyping for traits with a low heritability in order to make faster progress in pineapple breeding. This type of approach has enabled the pineapple genetic improvement program to speed up the development process and reduce costs, thus making the program more attractive for industry participation.

The program will be able to deliver the desired efficiency and effectiveness gains through use of marker-assisted selection and Genome-wide Association Studies (GWAS). Marker assisted selection requires knowledge of markers highly associated with traits of interest. Once associated loci are identified it is possible to propose candidate genes and proceed with further genetic analysis. It is also possible to develop reliable low-cost markers, such as KASP markers, for traits of interest that can be used to screen selected varieties.

Genomic selection is a relatively new approach in fruit breeding. This approach relies on a high throughput sequencing platform to generate thousands of markers across the genome. A ‘trainer population’ of seedlings is phenotyped and genotyped and the data analyzed to estimate genomic breeding values for each variety. The ‘trained’ models can then be applied to genotypic data from plants at a very early stage of plant growth/ development to estimate their traits or their breeding value. A genomic wide approach such as this is particularly useful where selection is for complex, polygenic traits. A genomic selection approach can be effective for all traits of interest simultaneously.

Marker-assisted selection and genomic selection have not been used previously in pineapple and hence the approach needed to be developed by this project.

As well as develop new varieties, the project aimed to develop a genomic wide marker

system and genomic selection protocol for improved fresh fruit pineapple varieties, an approach which is essential to speed up breeding of pineapple and reduce costs for industry.

Methodology

Breeding strategy

The current industry varieties 73-50 and 73-114, Hawaiian in origin, are the most successful modern varieties and are close to the desired type but with some major flaws. It is reasonable to assume that further breeding should stay close to these genotypes. Small but significant improvements are required without rearranging the genome excessively.

Experience so far indicates that a low level of inbreeding (half-sib and parental backcrosses) is feasible in pineapples and can avoid significant self-incompatibility (SI) and excessive inbreeding depression. This approach should produce populations slightly more homozygous for desirable traits, thus minimizing heterozygosity, the single major impediment to rapid breeding success in pineapple. This program followed this breeding strategy by utilizing a low level of in-breeding but with some outcrossing to ensure hybrid vigor was captured within the breeding populations.

The program utilized varieties developed in previous programs and the successful Hawaiian varieties MD-2 and its half-sib 75-970. The program followed a mild inbreeding strategy to avoid excessive heterozygosity and the associated introduction of undesirable traits. Some mild outcrossing to non-foundation varieties was used to recapture additive genetic effects. There was an emphasis on the use of the variety MD-2 to improve vitamin C content, blackheart resistance and reduce peduncle length. A range of other varieties developed from 73-50 were used as well as several Hawaiian varieties (PRI 62-85 and PRI 59-656) with putative resistance to phytophthora root disease.

A population of 35,000 seedlings was developed and planted on industry farms throughout Queensland.

Main traits

- High plant crop marketable yield (Greater than 73-50/MD-2). Fruit weight minus crown of 1.6-2.0 kg at 60,000 plants/ha. Size count 9-10 per carton.
- One vigorous sucker at harvest in summer
- Few slips in summer (<2)
- Short peduncle (Similar to MD-2)
- Smooth leaves
- Large crown (>200g) to ensure rapid plant establishment
- Moderate to good resistant to *Phytophthora cinnamomi* (Pc)
- Moderate to good resistant to natural flowering (NI) (similar to Cayenne or better)
- Good resistance to translucency

- Good resistance to internal browning (IB)
- Good resistance to fruitlet core rot (FCR)
- Yellow flesh (=MD-2/73-50)
- Mod-high TSS (winter >15%; summer >16%)
- Low acidity (0.4% in summer: <0.7% in winter)
- Low to moderate fibre (=MD-2/73-50)
- Low porosity
- Small to moderate core
- Pleasant flavor profile

The predominant recurrent parents were MD-2, its' half-sib PRI 75-970 and some DAF selections including Aus-Festival. The putative *Pc* resistant varieties Aus-Carnival (including offspring); Aus-Jubilee (including offspring), PRI 62-85, PRI 59-656, P2-6424, P1-1470, 58-729 and 21-1605 were used as parents to potentially impart resistance to *Pc*. The putative NI resistant varieties Aus-Carnival, Aus-Jubilee, 56-1401, 1-13134, P1-6424, P2-10523, P1-4152 and P1-45 were used as parents to impart resistance to NI. A small number of other parent combinations were used to incorporate high TSS, low fibre, red skin or different aromatic flavor compounds.

The majority of the selections were decided by the project leader in consultation with the grower co-operators on whose farms the trials are situated. Usually, a representative group of fruit from each harvest was evaluated with the grower on site. In the northern Townsville site, the project leader collaborated with the grower to have fruit harvested and evaluated in his absence.

Genomic Selection

The first step in developing a genomic selection protocol was to develop and analyze a training population of diverse genotypes that represent the breeding pool. A trial of 840 plants, comprised of 512 seedlings of 17 families and 53 parental genotypes was harvested over Jan-February, 2022. Data was collected on fruit and plant attributes and frozen pulp samples were retained for later analysis. The samples were sequenced by Diversity Arrays Pty Ltd and a preliminary Genome-wide Association Study (GWAS) analysis was completed on 23 traits of interest. The data set was randomly assigned into training and test populations and genomic selection analysis conducted for two traits to gauge the utility of the process. A Marker-Assisted BLUP function in Gapit 3.0 was used.

Molecular marker validation and application

The study included a broad genetic background of germplasm to ensure reliability. A suite of 48 KASP markers covering at least 20 traits was used to characterize a panel of genotypes with some known traits to allow validation. Trait and genotype data from the Genomic Selection trial was used. The KASP markers are currently being assayed by an industry provider, GeneWorks.

Potted experiments for NI in response to winter and drought stress

This experiment was reported on in a previous project but is included to give continuity

to the marker development component. Here the genomic data was re-referenced and re-analyzed using updated algorithms.

Approximately 1,300 seedlings of the susceptible variety MD-2 (♀) and more resistant variety Smooth Cayenne (♂) were obtained by cross pollination. The plants were grown for 9 months in 150 mm pots and by that time were considered large and mature enough to flower.

A trial to identify genotypes most susceptible to NI was completed over the winter and spring periods of 2016. In this trial, 400 large potted seedlings were exposed to ambient winter conditions in full sun and 274 were placed in a heavily shaded enclosure (67% shade) immediately adjacent to the full sun plants. Following this, those plants that did not flower were exposed to a second winter. Seventy-five seedlings that flowered from either the sun or shade groups were selected as susceptible. Those that did not flower after 2 winters were considered as resistant.

170 of those that did not flower following the first winter were removed from their pots without disturbing the root system and left on a bench in an open area for 2 days to dry the root system then repotted and watered. The plants were fertilized and watered until flowering was observed. Those that had flowered were considered susceptible to NI in response to drought stress.

Single Nucleotide Polymorphism (SNP) markers for each plant were re-analysed by genome-wide association study methods (GWAS) using the resistant and susceptible groups to identify the markers mostly reliably highly associated with resistance or susceptibility. The markers were analysed using the multi-loci mixed linear models BLINK and FarmCPU in the genome association prediction integrated tool (GAPIT 3.0), and the logistic regression, Generalised linear Mixed Model Association Test (GMMAT) algorithm.

Field Experiment for NI in response to winter

Initially, these plants were assessed for reaction to NI in the potted experiment. The majority of these potted plants were then multiplied over a 4-year period and a replicated field trial established. The trial contained 1,140 plants with 387 genotypes each replicated 1 to 4 times. This trial was planted so it would be challenged for NI response over the winter of 2022. SNP markers for each plant were analysed by GWAS as per the potted plant experiment.

Development of Molecular Markers for Resistance to Pc

This work was reported previously. The data here was however re-analysed by GWAS using different algorithms as for NI studies and is reported here.

Screening of germplasm for resistance to *Dickyea* strains

This work involved 2 separate trials. In the first trial the Mareeba strain of Dickeya was used. In the second trial both the Mareeba and NT strains were used. On both occasions, a very broad range of germplasm, including wild types as well as advanced DAF fresh fruit varieties were screened to explore potential genetic resistance. The Northern Territory strain is considered closest genetically to the highly virulent Malaysian strain. Because of the biosecurity concerns, this isolate could not be brought into Queensland and hence the screening was done in collaboration with the Northern Territory Dept Industry, Tourism and Trade at Berrimah, Northern Territory. The second trial is of most interest. In this trial, 17 pineapple accessions from wild to modern domestic were inoculated with the isolate *Dickeya zeae* NT-43531. The

inoculations were performed at the Berrimah Research Station. Data was collected on 3 occasions, 15/9/2021, 22/9/2021 and 24/9/2021. Data from each assessment was analyzed separately using analysis of variance (ANOVA), as well as combined into a repeated measures ANOVA.

Germplasm maintenance

The Australian pineapple germplasm collection was maintained each year and expanded with new varieties with promising future application.

High-throughput Sequencing, Variant Identification and Gene Annotation

DNA was extracted from pineapple leaf samples using silica column-based methods. Libraries were prepared from the samples using an Illumina DNA Prep (M) tagmentation kit and 150bp, paired-end sequencing performed using the Illumina NovaSeq S4 platform.

Paired-end reads were aligned using BWA to genomes from pineapple varieties F153 (GCF_001540865.1 and GCA_902162155.2), MD2 (ACMD2v2P0) and bracteatus CB5 (Acb.HiCasm20190122). Subsequently, variant identification (SNPs and indels) and joint genotyping was performed using GATK4. Gene effects of SNPs and indels were classified using SnpEff.

Available gene sequences for F153 (GCA_902162155.2) and MD2 (ACMD2v2P0) were BLAST-annotated using known *Ananas* proteins and the uniref50 database available from UniProt (www.uniprot.org). GO terms were retrieved using the core ontology dataset available from the Gene Ontology knowledgebase (www.geneontology.org). General feature format (gff) files for F153 and MD2 were subsequently annotated with BLAST scores, PFAM descriptions and GO terms using a custom Perl script. Annotations were visualized with Integrative Genomics Viewer (IGV).

KASP Marker Assays

Genomic regions flanking DArT markers and variants associated with numerous pineapple varieties were visualized using a JBrowse genome browser available at the Pineapple Genomics Database (www.pineapple.angiosperms.org). Custom Perl scripts were used to extract flanking genomic regions and convert variants to degenerate IUPAC codes for subsequent KASP assay design.

SNPs and indels identified in this study were visualized using Integrated Genomics Viewer (IGV). 100bp either side of each DArT marker of interest was retrieved from the F153 genome and variable bases converted to IUPAC codes manually. Sequences were then submitted to GeneWorks for KASP assay design using LGC's proprietary KBD software. Freeze-dried leaf tissue (0.05g) was submitted to GeneWorks/LGC in 96 well plates to perform the KASP assays.

Results and discussion

From the breeding component of the project over 330 preliminary selections were made across all sites. Collectively, the trial sites represented a very diverse set of growing environments. Cultural practices differed between trial sites and hence the seedling plants were challenged in many ways. In some cases, the trial plants generally performed poorly due to

droughted conditions. Excessive translucency was also observed on one trial site suggesting very low soil cation levels. In all cases, selections were made in comparison to industry standard varieties included in the trial or harvested nearby. Two fields were ratooned cropped and additional selections made on the ratoon fruit. In all harvests, selection pressure was applied for strong, vigorous, plants as well as fruit quality traits.

Preliminary data for the selections is included in appendix 1.

Genomic Selection

The Genomic Selection/ GWAS trial on Maroochy Research Facility was harvested over Jan-Feb, 2023. The Genomic Selection best linear unbiased prediction (BLUP) values for 5 varieties for two traits of interest are shown in table 1 along with the industry standard variety, 73-50.

Table 1. BLUPs and Prediction values for 2 traits on interest.

Trait	Variety	BLUP	PEV	BLUE	Prediction
Fruitlet Number per spiral	S_198	0.6679	0.544788	16.10475	16.8
	S_523	0.602155	0.458288	14.46568	15.1
	S_266	0.576651	0.449328	15.71085	16.3
	S_388	0.543738	0.584577	14.05595	14.6
	S_160	0.537022	0.476961	15.20754	15.7
	73-50	-0.01371	0.389228	12.62538	12.6
Fruitlet Diameter	S_58	0.877219	0.525968	22.9055	23.8
	S_405	0.74719	0.661613	20.91748	21.7
	S_357	0.658585	0.53589	21.48888	22.1
	S_135	0.623781	0.544227	23.34575	24.0
	S_376	0.582633	0.458469	21.36852	22.0
	73-50	0.076116	0.454187	22.94391	23.0

The two traits, fruitlet diameter and fruitlet number, included in table 1 are the main components of fruit size. An increase in either of these traits will give larger fruit size. From table 1, it can be seen that all varieties exhibited a higher BLUP (breeding value) with a higher trait value for fruitlet number and in one case, fruitlet diameter. This demonstrates Genomic Selection can be useful in pineapple breeding for improving fruit size using these individual traits.

Genome Wide Association Study (GWAS)

The GWAS results for a range of important plant and fruit traits were assessed.

Slip number is an important trait in pineapple. Generally, 1-2 slips are desirable as they can be a valuable source of planting material but a greater number exhausts plant assimilates and requires labour to remove them. An excessive number of slips is highly undesirable. Some of the SNPs identified were located in regions of the genome in close proximity to genes with functions that could associated with slip numbers and are also considered good candidates for further investigation.

Leaf size, comprised of the components area, length and width exhibited a relatively small number of associated markers. The best candidate SNPs were related to either morphogenesis

or photosynthesis genes. Several other SNPs were associated with leaf traits and candidate genes proposed.

The pineapple fruit is a compound structure comprised of many fruitlets arranged in spirals. Usually, there are 8 short spirals. Fruit size is a combination of fruitlet number and fruitlet size. Fruitlet number and fruitlet diameter were associated with 5 separate loci. Good candidates for these included some morphology-related genes and homeodomain genes but also included a number of other genes for berry size in fruit.

While it is expected that fruit length should be determined by similar loci as fruitlet number and fruitlet diameter, in the analysis here, different loci were identified. Similar candidate genes were however close to those loci.

The presence of knobs around the base of fruit is considered a highly undesirable trait. In this study, several loci were associated with the presence of basal knobs.

Crowns are the most important source of planting material. Larger crowns establish and grow faster than small crowns and hence are highly desirable. There were several loci found associated with crown weight.

Core diameter is an important consumer orientated trait, and it can be excessive in pineapple. Several associated loci were identified here but, while candidate genes were proposed, none are obvious.

Molecular markers for resistance to *Phytophthora cinnamomi*

A number of key genes were examined as potential molecular markers for resistance to *Phytophthora cinnamomi*. Of these genes, some included those associated transcription factors, stress and plant defence response and a gene well known to be associated with plant response to pathogen attack.

Drought Stress Flowering

Three SNP markers were found associated with NI from drought stress but only one was common to all algorithms and is considered to represent the main QTL. Of the genes around that marker, an ethylene pathway gene and a stress signaling gene are of interest.

The ethylene pathway is well-known to be involved in the flowering response in pineapple and stress in plants. The pathway involves both biosynthesis and signaling pathways so many genes are potentially involved.

Winter Initiated Flowering

From the potted and field trials involving NI from winter, 6 SNP markers were found associated. Nearby genes included those associated transcription factors and in plant morphology and flowering.

Several genes well-known to be involved in flowering in many species are also close to associated markers.

Dickeya test

Two trials assessing a subset of genotypes for resistance to *Dickeya* was completed. In the first test, the Mareeba isolate (BRIP 65197) was used. The area under the disease progress curve (AUDPC) was calculated (Table 2). Results indicate there is a significant difference between the accessions ($p<0.001$). The largest mean AUDPC (area under the disease progression curve) and hence greatest disease progression were for MD-2 and 73-50. Both of these accessions had mean AUDPC significantly greater than all other means.

Table 2. AUDPC data for genotypes inoculated with *Dickeya*.

Line	Accession	AUDPC	
1	Queen	2.063	c
2	Cayenne	1.670	c
3	Tapricanga	2.063	c
4	73-50	2.863	b
5	MD-2	3.820	a
6	Perola	1.670	c
7	Ruby	1.813	c
8	59-656	1.670	c
9	FRF 223	1.670	c
	p-value	<0.001	
	se	0.1998	
	SED	0.2826	
	95% LSD	0.5937	

Table 3 shows lesion progression over time. Comparisons are again made down each

column. When the mean disease severity ratings are compared for each accession over time, there is no significant change over time for accessions 59-656, Cayenne, FRF223, Perola, Ruby and Tapiricanga. Accession 73-50 showed a significant increase at the start of the trial while Queen increased significantly towards the end of the trial. MD-2 was the only accession which showed a significant and steady increase over time.

Table 3. Disease lesion progression over time (mm).

Day	59-656	73-50	Cay**	FRF223	MD-2	Perola	Queen	Ruby	Tapir*
2	41 a	42 b	43 a	38 a	39 d	38 a	40 b	39 a	51 a
4	41 a	75 a	43 a	38 a	65 c	38 a	40 b	39 a	51 a
7	41 a	73 a	43 a	38 a	70 c	38 a	40 b	39 a	51 a
9	41 a	73 a	43 a	38 a	88 b	38 a	40 b	43 a	51 a
11	41 a	75 a	43 a	38 a	92 a	38 a	40 b	43 a	51 a
14	41 a	75 a	43 a	38 a	99 a	38 a	51 a	43 a	51 a

NB: 59-656 is resistant to *Phytophthora cinnamomi*; FRF223 is a wild microstachys. Tapir* is Tapiricanga. Cay** is Cayenne

MD-2 and 73-50 are the main fresh market varieties in Australia. Both varieties appear more susceptible to the Mareeba isolate than other genotypes.

In the second test (table 4), seventeen pineapple accessions were inoculated with the isolate *Dickeya zeae* NT-43531. This isolate is considered very close to the highly virulent overseas pathogen. The inoculations were performed at the Berrimah Research Station, Northern Territory. Data was collected over time on 3 occasions, 15/9/2021, 22/9/2021 and 24/9/2021. Data from each assessment was analyzed separately using analysis of variance (ANOVA), as well as combined into a repeated measures ANOVA.

Manzana, CO42, Ruby and Red Spanish had smaller lesions than all other genotypes. The wild accession FRF223, and the 2 DAF varieties Aus-Festival and Aus-Jubilee had the largest lesions.

In all, 4 isolates have not been tested over 3 trials. The data from these trials suggests, while there is some tolerance to *Dickeya*, strong resistance is unlikely to exist. Higher levels of tolerance appear to exist in the genotypes Manzana (*syn*. CO42), Ruby and Red Spanish which have not been used to any extent in the current breeding program due to other undesirable traits. Two of the DAF selections appear to be highly susceptible to *Dickeya*. The DAF variety, Aus-Carnival (AC), while not represented in all tests, did display some tolerance where tested.

The general trend for all tests conducted appears to be that wild, primitive genotypes are the most susceptible.

Table 4. Mean lesion length data for 3 assessments over 9 days using the isolate *Dickeya zeae* NT-43531.

Accession	Mean lesion length (mm)	Signif. P<0.001	Rank (higher is longer lesion)
MD-2	32.92	e	9
Aus-Jubilee	56.88	ab	16
Aus-Festival	50.83	abc	15
Ruby	19.58	gh	2
Perola	32.08	e	8
FRF223	60.21	a	17
FRF414	46.67	bc	14
Manzana	17.29	h	1
Perolera	31.25	ef	7
McGregor	28.96	efg	6
White	45.42	cd	13
Red Spanish	20.62	fgh	3
73-50	35.21	de	12
59-656	33.33	e	10
Cayenne	28.46	efg	5
CO42	26.46	efgh	4
Tapiricanga	33.96	e	11
p-value	*<0.001*		
se	*3.722*		
SED	*5.264*		
95% LSD	*11.160*		

High-throughput Sequencing, Variant Identification and Gene Annotation

Using 150bp paired-end DNA reads, we obtained greater than 20x genome coverage for 31 pineapple varieties with an average coverage of 35x (Table 5) and a maximum coverage of 90x for cultivar 73-50. Coverage levels above 20x are generally considered sufficient to capture the majority of genetic information within a variety. Subsequent analysis of SNPs and indels within these 31 varieties revealed high variant frequencies, in accordance with our selection of a genetically diverse range of cultivars and high heterozygosity of pineapple genomes (Table 6). SnpEff was used to classify and annotate these variants according to their predicted gene effects (eg. protein truncation, mistranslation, synonymous mutation, etc) and revealed an unusually high number of effects within the original version of the F153 reference genome, for reasons that remain to be determined.

It is anticipated that sequence and variant information obtained during this project will form a valuable resource for a range of breeding applications within the Australian pineapple industry. Association of genetic variants with gene effects and the development of software to functionally annotate genes within this project provides useful tools to pinpoint causal mutations associated with traits of interest. Phenotyping each of the 31 sequenced varieties for a range of

traits would open up valuable opportunities for applications such as GWAS and genomic selection that can complement analyses conducted with DArT derived data in the current project.

KASP Assays

Kompetitive allele specific polymerase chain reaction (KASP) assays were identified as a potential method to cost-effectively validate markers identified by DArT and GWAS and subsequently genotype large, diverse populations. Effective KASP assay design depends upon the identification of conserved DNA regions flanking the marker of interest. The online Pineapple Genomics Database initially allowed us to visualize flanking DNA regions in several pineapple reference genomes and examine their conservation in a diverse range of cultivars. Subsequently, we wrote software to extract these regions and summarise all nucleotide variants within them for KASP assay design. Unfortunately, this online resource was suddenly closed before this process could be completed.

In order to complete KASP assay design, we subsequently utilized SNP and indel information obtained for the 31 varieties sequenced within this project. Sequence and variant data was used to cross-validate the DArT markers and identify conserved flanking regions for KASP assay design. Using this approach, 42 KASP genotyping assays were successfully designed for 41 SNP markers associated with 20 different traits (Table 7). The KASP assays were configured to work across all 31 varieties used in their design and are anticipated to possess broadly utility across a diverse range of pineapple cultivars. Sequences for an additional 6 assays are currently being processed by the provider in order to generate 48 KASP assays in total.

380 pineapple varieties will be screened with a total of 48 KASP assays. These varieties include 260 fresh fruit selections from PI17000, 39 additional varieties with putative resistance or susceptibility to stress-induced NI, 40 varieties with putative resistance or susceptibility to winter-induced NI and 40 varieties with putative resistance or susceptibility to root disease. Resistant and susceptible varieties were balanced for each trait. This will allow validation of some of the key markers as well as provide information on many of the preliminary selections developed in PI17000.

The identification of many of the markers shown in Table 7 occurred under project AS19003 and will be reported there. Traits were chosen for KASP assay based on the likely usefulness of the markers. A small number of easily phenotyped traits, such as leaf margin type, were included to test the reliability of the KASP marker system.

Table 5. Varieties used for KASP assay design and the coverage obtained (average coverage for Illumina 150bp paired end reads).

Variety	Average coverage (x) *
10_1985	39.65
10_1560	37.16
21_1605	37.85
55_2963	45.27
Aus_Festival	43.74
Aus_Jubilee	24.08
PRI_75_970	44.19
Red_Spanish	26.22

Smooth Cayenne (clone F97_626)	35.39
Selvagem_6	24.79
Ruby	29.18
P2_4171	26.78
PRI_58_1184	30.28
PRI_59_443	31.40
PRI_59_656	32.88
Primavera	30.26
P1_1470	33.09
58_729	40.880
P1_6424	36.86
56_1401	31.44
73-50	89.72
Tapiricanga	29.25
Manzana	25.87
Aus_Carnival	25.90
P2_3309	33.42
58_1137	28.40
FRF_414	35.99
MD2	44.71
Monte_Lirio	34.38
PRI_62-85	50.55
PRI_71_113	28.48
* Calculated using the MD2 reference genome.	

Table 6. Variant frequencies and associated gene effects for 31 joint-genotyped pineapple varieties.

Reference Genome	Variant type	Number of Raw Variants *	Raw Variant rate * (1 variant per X bases)	Number of Gene Effects *
MD2v2	SNP	24,592,613	22	38,363,982
	indel	3,924,158	146	6,897,196
F153v1	SNP	22,130,068	17	83,930,758
	indel	3,842,191	105	17,083,789
F153v2	SNP	22,131,213	17	35,929,825
	indel	3,843,225	105	6,482,683
CB5	SNP	27,870,073	18	43,346,009
	indel	5,284,115	102	9,157,827
* Values are combined across all 31 varieties.				

Table 7. Traits chosen for KASP marker assay.

Trait	SNP marker
Fruitlet Number	54313218
Fruitlet Number	4714774
Fruitlet Number	28880948
Fruitlet Diameter	28877593
Fruitlet Diameter	28880729
NI Winter Pots	28876150
NI Winter Pots	4717892
NI Winter Field	28877417
NI Winter Field	54313695
Fibre	4714126
Flavour	54307772
Flavour	4714421
Translucency	4727983
Translucency	4722716
Acidity	4711197
Acidity	28873623
TSS	54308139
TSS	4715563
TSS	28873947
TSS	54308299
Total Sugars	28882252
Total Sugars	54312300
Sucrose g/100g	4716713
Sucrose g/100g	54311560
Sucrose g/100g	28881610
Sucrose g/100g	54312039
Organic Acids	28873796
Malic Acid g/100g	54310098
Knobs	4714539
Phytophthora	100046224
Phytophthora	100067858
Phytophthora	4717643
Leaf Width	54309416
Core Diameter	4713938
Core Diameter	4709786
Core Diameter	28878349
Flesh colour Cielab b	54308994
Flesh colour Rating	28879211
Flesh colour Rating	4713595
Flesh colour Rating	4713595
Flesh colour Rating	4714993
Spiny leaf margin	4714214

Outputs

Table 8. Output summary

Output	Description	Detail
>330 preliminary selections	All selected plants are held on MRF pending further multiplication. The plants held include crowns, slips, suckers and stem sections. Plant number vary from 1 to approx 10 per variety.	The current project MRT for PI2201 outlines the next steps in multiplication and evaluation. This proposal includes testing in all growing regions thereby ensuring wide industry engagement.
Molecular markers identified for many important traits including resistance to NI and Pc. This includes identification of putative resistance genes.	An additional field trial looking at NI resistance was completed with some genotypes replicated. This was conducted over the winter of 2022 when the incidence of NI was very high across most pineapple growing regions.	As a first step, KASP markers for resistance to Pc and NI are currently being developed for application in variety selection.
Genomic Selection protocols established for several traits.	Two traits were tested.	This was achieved using the R platform program GAPIT 3 with the marker-assisted BLUP algorithm, MABLUP.
Resistance trends to *Dickeya* sp. across a range of genotypes.	No usable strong resistance was identified other than the current industry varieties MD-2 and 73-50 appear more susceptible than all other genotypes tested.	This was achieved over 2 separate experiments using 2 isolate of *Dickeya* sp. A very broad range of genotypes from wild to domestic was used to ensure capture of any potential resistance genes.
Australian pineapple germplasm collection maintained.	This collection is held at Maroochy Research Facility.	The collection is split over at least 2 plantings per year depending on the planting material availability. Occasionally varieties are multiplied to ensure they are not lost.

Outcomes

Table 9. Outcome summary

Outcome	Alignment to fund outcome, strategy and KPI	Description	Evidence
The pineapple now has future options for managing natural flowering and Root Rot that do not include heavy chemical use or compromise on planting schedules.	Outcome 2: Industry supply, productivity and sustainability. Strategy: Develop new varieties with superior production and disease resistance and enhanced consumer appeal.	Putative candidate genes for resistance to Pc and resistance to NI from drought and winter proposed. SNP markers significantly associated identified.	Trial data analysis. One peer-reviewed paper published on Pc resistance and one presentation given (IHC 2014).
Future options for varieties with improved production and quality attributes that are better suited to each region are on hand.	Outcome 2: Industry supply, productivity and sustainability. Strategy: Develop new varieties with superior production and disease resistance and enhanced consumer appeal.	Over 330 preliminary selections are held on MRF undergoing further multiplication.	Initial seedling selection trials were conducted on grower properties, often with growers.
Improved breeding and selection protocols for pineapple now developed.	Outcome 2: Industry supply, productivity and sustainability. Strategy: Develop new varieties with superior production and disease resistance and enhanced consumer appeal.	Most of germplasm collection now sequenced and referenced against latest genomes. Genomic Selection and GWAS completed for most traits with good associations.	Trial data for Genomic Selection analyzed for 2 fruit traits contributing to fruit size. GWAS completed for many traits. A trait with known gene control (leaf margin) included as a guide to model accuracy.

Monitoring and evaluation

Table 10. Key Evaluation Questions

Key Evaluation Question	Project performance	Continuous improvement opportunities
To what extent has the project achieved its expected outcomes?	The project aimed to develop >35,000 seedlings and evaluate them across all growing districts to select >100 individuals. All targets were achieved with >330 selections made.	Opportunity exists to improve parental selection using genomic tools developed under this project. A large collection of potentially valuable parental lines is now on hand.
How relevant was the project to the needs of intended beneficiaries?	The majority of the selections were done on grower properties across all major growing regions. Usually, the evaluation and selection were done with the grower, the main immediate beneficiary, ensuring their needs are included.	More growers can be involved in the process including the younger generation of growers.
How well have intended beneficiaries been engaged in the project?	In most cases it was mainly the co-operating growers that were involved although several talks were given to grower groups. This is the only practical option at this early stage of the program.	As selections are multiplied and planted in growing regions, opportunity exists to have more growers involved in evaluating the varieties.
To what extent were engagement processes appropriate to the target audience/s of the project?	In most cases, the co-operating growers were industry leaders with successful farming businesses. The broader industry was engaged through the annual field day which is attended by most growers and through regional study groups.	Further grower involvement will be achieved when there are sufficient plant numbers of preliminary selections.
What efforts did the project make to improve efficiency?	The project embarked on an ambitious objective of gathering data on genetic determinants of the main traits of interest. This included developing Genomic Selection protocols and identifying loci associated with each trait. In this way, candidate genes have been proposed for many traits in a world first result.	Genomic and marker datasets can be improved using more detailed platforms as these become cheaper and more available.
What efforts did the project make to maintain the Australian pineapple germplasm collection?	The germplasm collection was replanted at least twice over the life of the project with some genotypes multiplied to maintain their numbers.	Larger plots are needed for most of the varieties to ensure they are not lost. Disease control, mostly *Pc*, needs to be improved.

Recommendations

PI17000 has developed a large number of preliminary selections with good eating quality and yield and potential improved resistance to Root Rot and Natural Flowering. The involvement of a geographical spread of co-operating properties which encompassed different growing environments and cultural practices revealed a large difference in associated fruit quality specific to those sites. So, while new varieties have been developed, the production of varieties with improved quality also requires growing systems to be optimized if those quality improvements are to be realized. It is highly recommended the industry revisit soil health and nutritional practices with an improved focus on cation availability. Failure to optimize cultural practices limits potential benefits from new varieties.

It is also recommended that harvest maturities and supply chain practices be optimized with a focus on consumer preferences. This might include options for improved varietal branding.

Refereed scientific publications

Journal articles

Sanewski G. 2022. DArTseq molecular markers associated with the piping leaf margin phenotype in pineapple (*Ananas comosus* L.) Tropical Plant Biology.

Chen, LY., VanBuren, R., Paris, M. Sanewski, GM *et al.* The *bracteatus* pineapple genome and domestication of clonally propagated crops. *Nat Genet* **51**, 1549–1558 (2019). https://doi.org/10.1038/s41588-019-0506-8

Whole book

The Pineapple. Botany, Production and Uses (second edition). 2018. Sanewski G, Bartholomew DP and Paull RE. (Eds), CABI International, Oxfordshire.

Chapter in a book or paper in conference proceedings

Sanewski, G. 2018. The History of Pineapple Improvement. In: Ming R. (Ed), Genetics and Genomics of Pineapple. Springer Nature Switzerland, pp. 87-96.

Sanewski GM, Coppens d'Eeckenbrugge G, and Junghans DT (2018) Varieties and Breeding. In: The Pineapple. Botany, Production and Uses (second edition). Sanewski G, Bartholomew DP and Paull RE. (Eds), CABI International, Oxfordshire.

Jimenez JV, Sanewski GM, Reinhardt DH and Bartholomew DP. (2018) Cultural System. In: The Pineapple. Botany, Production and Uses (second edition). Sanewski G, Bartholomew DP and Paull RE. (Eds), CABI International, Oxfordshire.

Bartholomew DP an Sanewski GM (2018). Inflorescence and Fruit Development and Yield. In: The Pineapple. Botany, Production and Uses (second edition). Sanewski G, Bartholomew DP and Paull RE. (Eds), CABI International, Oxfordshire.

References

Allardyce JA, Rookes JE, Hussain HI and Cahill DM (2013) Transcriptional profiling of *Zea mays* roots reveals roles for jasmonic acid and terpenoids in resistance against *Phytophthora cinnamomi*. Functional Integrated Genomics. 13, 217-228.

Chunthawodtiporn J, Hill T et al. (2018) Quantitative trait loci controlling fruit size and other horticultural traits in bell pepper (*Capsicum annum*). The Plant Genome. March.

Feng et al. (2012) Identification of a novel nuclear-localized adenylate kinase 6 from Arabidopsis thaliana as an essential stem growth factor. Plant Physiology and Biochemistry. 61, 180-186.

Fletcher J. (2001) The ULTRAPETALA gene controls shoot and floral meristem size in *Arabidopsis*. Development. 128(8), 1323-33.

Gachomo E, Jimenez-Lopez JC et al. (2014) GIGANTUS1 (GTS1), a member of Transducin/ WD40 protein superfamily, controls seed germination, growth and biomass accumulation through ribosome-biogenesis protein interactions in *Arabidopsis thaliana*. BMC Plant Biology. 14, 37.

Huang Y, Du L et al. (2022) Multifaceted roles of zinc finger proteins in regulating various agronomic traits in rice. Front Plant Sci. 13, 974396.

Islam M, Rookes J and Cahill D. (2017) Active defence by an Australian native host, *Lomandra longifolia*, provides resistance against *Phytophthora cinnamomi*.

Jia et al. (2020) A serine/ threonine protein kinase encoding gene KERNEL NUMBER PER ROW6 regulates maize grain yield. Nature Communications. 11(988).

Kim et al. (1998) The ROTUNDIFOLIA3 gene of *Arabidopsis thaliana* encodes a new member of the cytochrome P450 family that is required for the regulated polar elongation of leaf cells. Genes Dev. 12(15), 2381-2391.

Koc E, Islek C and Kasko Arici Y. (2017) Spermine and its interaction with proline induce resistance to the root rot pathogen *Phytophthora capsici* in pepper (*Capsicum annum*). Horticulture, Environment and Biotechnology. 58(3), 254-267.

Kong W et al. (2019) Expansion and evolutionary patterns of glycolsyltransferase family 8 in Graminae crop genomes and their expression under salt and cold stresses in *Oryza sativa* ssp. Japonica. Biomolecules. 9(5).

Kou X, Liu C et al. (2016) NAC transcription factors play an important role in ethylene biosynthesis, reception and signaling of tomato fruit ripening. Molecular Genetics and Genomics 291, 2120-2217.

Lei Y, Sun Y et al. (2020) Woodland strawberry WRKY71 acts as a promoter of flowering via a transcriptional regulatory cascade. Horticultural Research. 7, 137.

Li J, Zhang Y et al. (2020) Plant U-box E3 ligases PUB25 and PUB26 control organ growth in *Arabidopsis*. New Phytologist. http://doi.org/10.1111/nph.16885.

Liu JF, Qigo YY et al. (2023) The NAC transcription factors play core roles in flowering and ripening fundamental to yield and quality. Frontiers in Plant Science. 23.

Munoz-Espanosa et al. (2020) Identification of SNPs and InDels associated with berry size in table grapes integrating genetic and transcriptomic approaches. BMC Plant Biology. 20(365).

Narita et al. (2004) Overexpression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in *Arabidopsis thaliana*. The Plant Journal. https://doi.org/10.1111/j.1365-313x.2004.02078.x

Quan C, Bai Z et al. (2019) Genome-wide analysis and environmental response profiling of phosphate-induced-1 family genes in rice (Oryza sativa). Biotechnology and biological Equipment. 33(1).

Regierer et al. (2002) Starch content and yield increase as a result of altering adenylate pools in transgenic plants. Nature Biotechnology. 20(12), 1258-1260.

Seifi HS, Zarei A, Hsiang T and Shelp BJ (2019) Spermine is a potent plant defense activator against gray mold disease on *Solanum lycopersicum*, *Phaseolus vulgaris* and *Arabidopsis thaliana*. Phytopathology. Epub April 16.

Sen S, Kundu S and Dutta SK (2016) Proteomic analysis of JAZ interacting proteins under methyl jasmonate treatment in finger millet. Plant Physiology and Biochemistry. 108, 79-89.

Slovak et al. (2020) Ribosome assembly factor Adenylate kinase 6 maintains cell proliferation and cell size homeostasis during root growth. New Phytologist. 225(5), 2064-2076.

Snouffer et al. (2020) The shape of things to come: ovate family proteins regulate plant organ shape. Current Opinion in Plant Biology. 53, 98-105.

Wang B et al. (2012) UGT87A2, an *Arabidopsis* glycolsyltransferase, regulates flowering time via Flowering Locus C. New Phytologist. 194(3), 666-675.

Wang et al. (2019) Krn1, a major quantitative trait locus for kernel row number in maize. New Phytologist. 223(3), 1634-1646.

Wang J, Wang H et al. (2022) The role of the NAC transcription factors in flower development in plants. Chinese Journal of Biotechnology. 38(8), 2687-2699.

Wang Y et al. (2012) Over-expression of a putative poplar glycosyltransferase gene, PtGt1 in tobacco increases lignin content and causes early flowering. J of Experimental Botany. 63(7), 2799-2808.

Yadu B, Chandrakar V, Meena RK et al (2018) Spermidine and melatonin attenuate fluoride toxicity by regulating gene expression in antioxidants in *Cajanus cajan* L. Journal of Plant Growth Regulation. 37, 1113-1126.

Yoo SY, Kim Y et al. (2007) Control of flowering time and cold response by a NAC domain protein in Arabidopsis. PLOS One 2(7).

Zhang G et al (2017) Ectopic expression of UGT84A2 delayed flowering by indole-3-butyric acid-mediated transcriptional repression of ARF6 and ARF8 genes in *Arabidopsis*. Plant Cell Reports 36(12). 1995-2006.

Zhang J et al. (2022) Transcriptomic analysis reveals potential gene regulatory networks under cold stress of loquat (*Eriobotrya japonica* Lindl.). Frontiers in Plant Science. 13.

Zhang Y, Ziwen L et al. (2020) Phylogeny and functions of LOB domain proteins in plants.

Int J Mol Sci. 21(7), 2278.

Zou X and Sun H. (2023) DOF transcription factors: Specific regulators of plant biological processes. Front Plant Sci. 14. http://doi.org/10.3389/fpls.2023.1044918.

Intellectual property

No commercialisable project IP or commercialization to report. All field plantings on industry farms were covered by standard DAF Material Transfer Agreements and growers recompensed for plant care.

Acknowledgements

I wish to acknowledge the support by the Australian Pineapple Industry and the individual growers that collaborated including Fullerton Farms, Pinata Farms, Forster Brothers Farming, Sandy Creek Pineapple Company, Wide Bay Pineapples, Lake Mary Pines and Pace Farming.