Final Report

New Genetic Targets to Improve Quality in Papaya

Project leader:

Rebecca Ford

Delivery partner:

Griffith University

Project code:

PP15000

Project:

New Genetic Targets to Improve Quality in Papaya – PP15000

Disclaimer:

Funding statement:

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

Publishing details:

ISBN 978 0 7341 4403 4

Published and distributed by: Hort Innovation

Level 8
1 Chifley Square
Sydney NSW 2000

Telephone: (02) 8295 2300

www.horticulture.com.au

Contents

Summary

1. Project objective and aims

The overall objective of the project was to improve the quality of Australian papaya to ensure consistency of supply to the consumer, and provide access to the new varieties, to improve growers' productivity and profitability. This was achieved through the following aims:

- Improve eating quality especially in flesh flavour of both yellow and red papaya.
- Tailor-make varieties suitable to the main growing regions (Tableland and Coastal areas)
- Replicate the new elite lines through seed production (priority) and micropropagation.
- Generate a database of material and understand the performance of exotic genetics in different growing areas to inform potential future selection of desirable characteristics to transfer to the Australian lines.
- Establish breeding and segregating populations for desirable traits identified by the industry and consumers including a range of tree and fruit quality traits.
- Develop bioassays for robust, reliable and practical evaluation of agronomic and fruit traits.
- Study the inheritance of selected quality traits within the advanced breeding populations.
- Develop papaya molecular markers, create a dense genome framework map of papaya and identify markers closely associated with the genome regions governing traits of interest
- Develop papaya with resistance to PRSV-P by introgression of resistance genes from *Vasconcellea* species.
- Inform research priorities and disseminate research outcomes through Industry Liaison Officer consultation with industry members and through publication and media.

2. Target audience

- The improved breeding lines are targeted for use by researchers, growers, breeders, marketers and consumers.
- The germplasm collection is targeted for use by researchers and breeders.
- The molecular genotype profiles are targeted for use by researchers and for the protection of plant breeders' rights.
- The developed evaluation tools are targeted for use by breeders and researchers.
- The genome tools, to determine key genome sequences responsible for conditioning desirable traits, are targeted for use by breeders and researchers.

3. Project activities

This project consisted of five sub-projects as listed below;

Breeding to improve commercial Australian papaya: Multiple F3 advanced breeding lines of red and yellow papaya were selected based on tree performance and planted in five independent trial sites representative of the Tablelands and Coastal regions. F4 seed was produced.

Papaya germplasm collection: A range of papaya germplasm was collected, evaluated for tree and fruit characteristics and genotyped using short sequence repeat (SSR) molecular markers. Trees were planted in three trial sites representative of the Tablelands and Coastal regions.

Molecular assisted papaya breeding: Mapping populations ('RB2', 'Sunrise Solo', F1 and F2) were assessed for hybridity and heritability of desirable fruit quality traits. The quantitative trait locations (QTL) of the major sweetness genes and other traits were uncovered through whole genome mapping using a Genotyping-By-Sequencing (GBS) approach. Candidate genes were investigated for functional association with sweetness through quantitative PCR.

Resistance to PRSV-P: The cross of *C. papaya* '2.001' x BC4#113 (*V. parviflora* x *V. pubescens*), which is resistant to PRSV-P, was successfully embryo-rescued, micropropagated, maintained and tested for hybridity.

Industry development: To ensure the research program is aligned with industry and consumer requirement and to relay the latest research outcomes to all stakeholders, the IDO frequently and regularly liaised with growers as well as the steering group and wider papaya industry and marketing groups. She organized extension events throughout the project lifespan for information gathering and dissemination.

4. Key outputs

- More genetically stable breeding lines of F4 of red and yellow papaya have been developed.
- The Papaya Evaluation Handbook was developed and published.
- A collection of broad papaya genetics was established, assessed and reported within a database.
- Eight major QTL and associated molecular markers for flesh sweetness, fruit weight, length, width and firmness were located on the genetic map and key genes associated to flesh sweetness were identified.
- PRSV-P resistance was introgressed into a cultivated papaya genetic background.
- Papaya Field Days and scientific dissemination events and publications.

5. Key outcomes

- The developed advanced breeding lines outperform the industry standard varieties in many of the desired agronomic and fruit quality traits including setting fruit lower on the tree for ease of harvest, higher and consistent yields and sweeter and pleasantly flavoured fruit.
- The developed standard protocols for trait measurement and reporting have improved the reliability of selection and communication within the entire papaya supply chain and the handbook has been widely adopted and implemented throughout the industry.
- The diverse genetics captured and databased within the germplasm collection represent a resource for future genetic gain through introgression into the elite cultivars.
- The molecular genotype fingerprints developed for the cultivars and broader germplasm are pivotal for providing Plant Breeders' Rights protection and determining genetic relatedness.
- Knowledge on the locations of major genetic components for flesh sweetness, fruit firmness, flesh thickness, skin freckle, and fruit size has been gained. This is the foundation for the development of molecular trait selection tools for breeders, saving time and costs associated with traditional evaluation and selection.
- The transfer of PSRV-P resistance to the cultigen is a major step forward in creating resistant superior varieties and the developed hybrid germplasm represents a foundation for this to occur.
- Knowledge related to papaya breeding, molecular breeding, bioinformatics and tissue culture of papaya has been transferred to papaya breeders, growers and researchers through scientific publications, industry communications and direct networking. The project has produced one PhD and one Honours graduate, and provided small research projects for five final year undergraduates. These represent upskilled individuals for the future growth of the industry.

- A trained industry development officer who is extremely familiar with all aspects of the Australian papaya industry and who has created multiple linkages with national and international papaya researchers. This has been the conduit in communication of relevant R&D outcomes to growers and supply chain partners via meetings, workshops, media and direct communication.

6. Recommendations

- The genetically stable advanced breeding lines should be further developed to be used as new cultivars and other parental lines to create new F1 hybrid cultivars.
- More genotypes should be sourced from overseas and within Australia to explore wider genetic resources. They may contain some useful characteristics, such as disease tolerances and can be stabilised to be used as parental lines for the development of new hybrids.
- Rapid/speed breeding techniques should be implemented in the breeding program together with the validation of genomics tools for accurate selection for traits including sweetness, flesh texture, skin quality and fruit size.
- Sweetness-related genes identified from Sunrise Solo should be assessed as suitable sweetness markers for possible stable transfer to other and newly developed elite advanced breeding lines. The candidate genes underpinning major QTL of other traits (fruit weight, fruit length, fruit width skin freckle, flesh thickness and fruit firmness characters) should be characterised and validated to accelerate the breeding program.
- The fruit sensory traits of flavour and taste should be deeply profiled to understand consumer preference and application in the advanced selective breeding.
- The newly developed advanced breeding lines should be assessed for disease response to common major pathogens to ensure that awareness exists regarding the potential requirement for disease management and with the optimal aim that future cultivars contain an acceptable level of resistance.
- The industry should continue to invest in the Industry Development Officer position who will continue to liaise with all stakeholders to ensure research is aligned exactly with industry needs and all are kept up-to-date with research outcomes.

Keywords

Breeding; germplasm; fruit quality traits; productivity traits; trait evaluation; fruit development; papaya genome; papaya ringspot virus

Introduction

Historical background to the project and significant linkages to industry strategy

Over the past seven years, the Griffith University-Horticulture Innovation Australia collaborative papaya breeding program (PP10005 and PP15000) has developed and validated a suit of highly accurate and robust breeding tools for the selection of a raft of targeted agronomic and fruit quality traits. The partnership has also developed advanced breeding materials for the future commercialization of superior red and yellow varieties. Together, we have made significant progress, backed up with scientific rigor and evidence, towards tailor-made varieties suitable to maximise yield and quality in each of the growing regions: The Tablelands and the Coastal areas. The advanced lines we have developed outperform the current industry standard varieties in yield, eating quality, flavor and several other desirable characters.

The initial phase of the papaya breeding program PP10005 (2011-2015) focused on germplasm evaluation and identification of new genetic resources. The major outcome was the establishment of segregating populations from selected outstanding lines with significant improvement of flavour, flesh texture, sweetness and other agronomic characters. These were the starting material for PP15000 (2015-2018), which has produced more stable genetic materials and standardised the evaluation methods to select for papaya fruit quality traits in multiple environments. Accordingly, genetically stable F4 lines have been established and evaluated to be superior for many target traits under multiple environments and in different cropping seasons. The further stabilisation and selection of individual lines will provide the industry with a brand-new set of elite cultivars that are entirely aligned with consumer, marketer and grower demands. These represent enormous potential to grow the Australian papaya industry and increase profitability.

Simultaneously, two other key traits were researched within the collaborative program. This included several sweetness-related genes being identified and characterised for use in marker-assisted selection of papaya sweetness. The F2 population for this was developed from a cross between RB2 and Sunrise Solo during PP10005. Subsequently, a PhD candidate (Miss Usana Nantawan) enrolled at Griffith University was aligned with PP15000 and identified the sweetness QTL and putative gene candidates. Also, for the last two decades, there has been a comprehensive effort to develop PRSV-P resistance in papaya, a major biotic constraint to expansion of the industry. Following on from initial discovery of resistance to the virus in *Vasconcellea pubescens*, in PP10005 and PP15000, this was transferred to papaya using *V. parviflora* as a bridging species, the first reported successful transfer of PRSV-P resistance from a *V.* pubescens to papaya. The result being that we now have resistance to PSRV-P within the cultigen and through careful backcrossing, this can be brought into elite cultivars.

All of the progress within the collaborative program has been underpinned by the Industry Development Officer, a position funded as a continuation from PP13007 into PP10005 and PP15000. This has resulted in clear communication of industry priorities to guide the research efforts and timely and accurate dissemination of research outputs back to all levels of industry. In summary, all outputs from the collaborative research programs were informed by and fed into the Papaya Strategic Investment Plan of "Increased quality, to ensure consistency of supply to the consumer, and provide access to the new varieties, to improve growers' productivity and profitability".

Methodology

1: Breeding to improve commercial papayas in Australia

1.1 Plant materials

Red papaya: Line R2-14, from PP10005, was selected and used as the starting material in this project. Two-hundred F3 populations were developed by single seed decent (SSD) and were evaluated three times for key traits (Table 1). The three best performing lines in each environment were selected in comparison with the industry standard variety (RB1) (Table 2). These were self-pollinated and advanced to F4. The trees were planted in December 2016 to February 2017 in multiple locations as listed below. Seed of F4 generations were collected and stored.

1. Lecker Farming, Mareeba (latitude -16.979592, longitude 145.329436)
2. Skybury Farmgate, Mareeba (latitude -17.010049, longitude 145.337257)
3. Joe Zappala, Innisfail (latitude -17.630151, longitude 145.963368)
4. Eddie Mizzi, Innisfail (latitude -17.598523, longitude 146.082866).

Yellow Papaw: Eleven lines of F2 yellow papaw containing the genetics of 1B, Honey Drew and TS2, which were developed in PP10005, were evaluated three times using the same method as the red papaya. These populations were planted at the Lecker Farming, Mareeba site. The F3 were collected and planted in March to April 2018 in three locations; Lecker Farming, Mareeba, Michael Oldano, Innisfail (latitude -17.663431, longitude 146.041650) and Jeff Schluter (latitude -17.598523, longitude 146.082866).

1.2 Seed and seedling preparation: Seed were extracted, cleaned and stored at 15ºC until further use. They were soaked in 2 mM gibberellic acid solution (Fermoz) for 15 minutes and then rinsed with tap water before sowing. A total of 120 seed were sown per generation at each location. This enabled growth of 50 sex-determined seedlings per location. Single seeds were sown using seed-raising mix (Searly, Australia) in 48-cell seedling trays. Trays were held in a shade-house, watered and treated according to the farm management practice at Lecker Farming to produce healthy seedlings for field-planting. Seedlings were sex determined at 4-6 weeks after germination and field-planted two months after sowing.

1.3 Sex determination at seedling stage: A leaf sample of each 4 to 6-week-old seedling was incubated in 50 mM sodium hydroxide and stabilised with 1mM Tris-Cl. The crude extract was used as the template for the PCR reaction. The method will be published in Acta Horticulturae 2018 (Appendix 1).

1.4 Evaluation of traits of interest: Forty trees of each generation at each location were evaluated three times (Table 1) as detailed in the Papaya Evaluation Handbook (Nantawan et al. 2017). The individual target traits within the advanced breeding lines were compared to those within the standard commercial varieties. The selection index was calculated by comparing each trait in the breeding line to the yellow or red industry standard variety (Table 2).

Table 1: List of traits assessed in the evaluation of papaya fruit quality and tree productivity.

No.	Tree age	Productivity traits	Fruit quality traits
1	5 months	Sex type Height to first marketable fruit Number of side shoots Trunk circumference Peduncle length	
2	10 months	Trunk circumference Peduncle length Saleable yield (Formula 1) Yield gap Number of carpelloid fruit Fruit ratio (formula 2)	Skin gloss Skin freckle Skin colour Fruit firmness Fruit shape Teat shape Stalk insertion Cavity shape Flesh colour Consistency in flesh colour Flesh thickness Flesh texture Flesh sweetness (ºBrix)
3	15 months	Same as the 2nd evaluation	Same as the 2nd evaluation

Saleable yield (kg) = Number of fruits per tree x Mean fruit weight (kg) ---------- **Formula 1**

$$\textbf{Flesh ratio} = 1 - \frac{\textbf{Cavity Width (cm) x Cavity Length (cm)}}{\text{Fruit Width (cm) x Fruit Length (cm)}} \quad \textbf{----------} \quad \textbf{Formula 2}$$

Table 2: Selection criteria and calculation of breeding index for tree productivity and fruit quality traits

1.Traits and their industry/consumer-derived preferred value or range		**2. Traits expected to have a higher value than in the industry standard**		**3. Traits expected to have a lower value than in the industry standard**	
Fruit weight:	Yellow: 0.9 – 1.5 kg Red: 0.9 – 1.2 kg	List of traits:	• Yield • Trunk circumference • Flesh ratio • Skin gloss • Skin colour • Fruit firmness • Flesh colour • Flesh thickness • Flesh texture • Flesh sweetness	List of traits:	• Height to first fruit • Number of side shoots • Number of carpelloid fruits • Number of wasted fruits • Skin freckle
Peduncle length:	Medium or long				
Fruit shape:	Yellow: Round shape #4 Red: Elongate shape #7				
Stalk insertion:	Flattened (#2)				
Teat shape:	Flat (#3)				
Cavity shape:	Round (#1)				
Flesh colour consistency:	>91%flesh colour consistency	Calculation of breeding index: Index = Value of breeding line / Value of the standard variety		Calculation of breeding index: Index = 2– Value of breeding line / Value of the standard variety	

2: Papaya germplasm collection

Thirty-nine lines of papaya were obtained from three sources: Australian commercial varieties, breeding lines and collections from our international collaborators (Table 3). Seed of 28 lines were sown in September 2016 from which only 17 lines germinated. These 17 lines and R3 (tissue culture lines) were DNA extracted and genotyped using 48 SSR primers (Appendix 2). An additional 11 lines were field planted and assessed in April 2018.

Seed were treated and sown as described above in section *1.2 Seed and seedling preparation*. They were field planted for evaluation at five locations; Lecker Farming, Mareeba, Skybury Farmgate, Mareeba, Joe Zappala, Innisfail, Michael Oldano, Innisfail and Jeff Schluter, Innisfail. Trees and fruit were assessed for specific traits following the protocols established within the Papaya Evaluation Handbook (Nantawan et al. 2017).

Table 3: Details of varieties for germplasm evaluation

Number	Code	Name	Descriptions	Note
1	**GR1**	RB1	Commercial red papaya in Australia	1
2	**GR2**	RB2	Commercial red papaya in Australia	1
3	**GR3**	RB4	Commercial red papaya in Australia	1
4	**GR4**	Solo	Commercial red papaya in Australia and Hawaii	1
5	**GR5**	Linda Solo	Red papaya	1
6	**GR6**	Holland	Commercial red papaya in Thailand	1
7	**GR7**	Khaeg Dum	Commercial red papaya in Thailand	N/A
8	**GR8**	Khaeg Nuan	Commercial green papaya in Thailand	1
9	**GR9**	First Lady	Red papaya in Taiwan	1
10	**GR10**	TN#2	Commercial green papaya in Taiwan	1
11	**GR11**	SKB-1	Breeding collection from Skybury	N/A
12	**GR12**	SKB-2	Breeding collection from Skybury	N/A
13	**GR13**	SKB-3	Breeding collection from Skybury	N/A
14	**GR14**	SKB-4	Breeding collection from Skybury	N/A
15	**GR15**	SKB-5	Breeding collection from Skybury	N/A
16	**GR16**	SKB-6	Breeding collection from Skybury	N/A
17	**GR17**	SKB-7	Breeding collection from Skybury	N/A
18	**GR18**	SKB-8	Breeding collection from Skybury	N/A
19	**GR19**	SKB-9	Breeding collection from Skybury	N/A
20	**GR20**	SKB-10	Breeding collection from Skybury	N/A
21	**GR21**	OP-Dimbulah	Open pollinated red papaya (Caliman) from Dimbulah	1
22	**GR22**	JC2	Breeding line from Prof Rod Drew	1
23	**GR23**	TS2	Breeding line from Prof Rod Drew	1
24	**GR24**	OP-Babinda	Open pollinated red papaya from Babinda	1
25	**GR25**	Paris	Open pollination of *C. papaya* cv. 'Paris'	1
26	**GY26**	H1B	Commercial yellow papaw in Australia	1
27	**GY27**	H11B	Commercial yellow papaw in Australia	1
28	**GY29**	H29	Commercial yellow papaw in Australia	1
29	**R1**	Red 24.2	Malaysian red, derived from tissue culture	2
30	**R2**	Solo 30.9	Solo Linda, derived from tissue culture	2
31	**R3**	Red #3	Skybury red, derived from tissue culture	1
32	**R4**	Solo 8.2	Tableland Solo, derived from tissue culture	2
33	**R5**	Solo 29.9	Sunrise Solo, derived from tissue culture	2
34	**R6**	F1 (27.1)	F1 self pollination of RB2	2
35	**R7**	F1 (30.9)	F1 self pollination of Solo Linda	2
36	**R9**	Pluk Mai Lai	Thai red papaya	2
37	**R10**	Holland #1	Thai red papaya	2
38	**Y15**	H13	Commercial yellow papaw in Australia	2
39	**HD1**	Honey drew 2.10	Breeding line of yellow papaya from PP10005	2
40	**HD2**	Honey Drew 2.11	Breeding line of yellow papaya from PP10005	2

Note: 1 = Trees planted in 2016, 2 = trees planted in 2018 and N/A = seeds failed to germinate

For developing molecular fingerprints, the DNA of five Australian commercial varieties (RB1, RB2, RB4, 1B and Sunrise Solo), a yellow parental line (ER62) and a Thai cultivar (Holland aka Sekaki) were extracted. A total of 45 individuals were genotyped to assess genetic diversity among and within cultivars using a Genotyping-By-Sequencing approach to detect single nucleotide polymorphisms (SNPs). The GBS service was carried out at the Australian Genome Research Facility (AGRF; Melbourne, Australia). Sequences were assessed for quality of the data, mapped to the *C. papaya* reference genome, 'SunUp' variant (Ming et al. 2008), called for variants and analysed for genetic diversity (Appendix 6).

3: Towards the development of a molecular-assisted program within the papaya breeding program

Detailed methods are provided in Appendix 3 (Nantawan's Ph.D. thesis).

3.1 Bioassay development: Five fruit maturity stages were studied to standardize the harvesting index and 22 characteristics (nine productivity traits: sex type, height to first marketable fruit, number of side shoots, trunk circumference, peduncle length, saleable yield, yield gap, number of carpelloid fruit, fruit ratio and 13 fruit quality traits: skin gloss, skin freckle, skin colour, fruit firmness, fruit shape, teat shape, stalk insertion, cavity shape , flesh colour, consistency in flesh colour, flesh thickness, flesh texture, flesh sweetness (ºBrix)) were studied in detail and standardised methods for evaluation and reporting were developed. These are summarised in the Papaya Evaluation Handbook (Table 1, Table 2, Formula 1, Formula 2; Nantawan et al. 2017).

3.2 Genetic inheritance: Three generations of the segregating populations derived from the cross of 'RB2' x 'Sunrise Solo' (Parents, F_1 and F_2) were assessed for fruit quality attributes and to determine the heritability of each of the traits. This was done with three replications and using the developed bioassays from the previous section.

3.3 Identification of candidate genes related to flesh sweetness: A combined and comparative genomics approach was undertaken, using DNA sequencing and qPCR, to uncover the putative sweetness-related genes within the papaya genotype 'Sunrise Solo'. Orthologous gene information for the candidate genes related to sweetness was sourced from The Arabidopsis Information Resource (TAIR, http://www.arabidopsis.org/) and NCBI database. Published gene sequences were blasted against aligned sequence of 'RB2' and 'Sunrise Solo' to locate genomic regions of candidate genes. A sequence corresponding to each published gene was compared between the two parental lines to uncover polymorphisms and primer pairs were designed accordingly to detect them. The differential expressions (transcriptions) of the genes of interest were compared between sweet and unsweet lines in nine tissue types (leaf, fruit at four maturity stages and fruit at four ripening stages) as further evidence of functional association with the sweetness trait.

3.4 QTL analysis of fruit quality traits: Flesh sweetness and other fruit quality traits (fruit weight, fruit length, fruit width, skin freckle, flesh thickness and fruit firmness) were subjected to quantitative trait loci (QTL) analysis using 226 F2 derived from the cross 'RB2' x 'Sunrise Solo'. High quality DNA was extracted from all samples. A GBS approach was used to detect single nucleotide polymorphisms (SNPs) between the parents and segregating in the F2 genomes. The GBS service was carried out at the Australian Genome Research Facility (AGRF; Melbourne, Australia). High quality SNPs were filtered using R script. They were subsequently used to construct a linkage map of the RB2 x Sunrise Solo genome. QTL analysis was performed with the genotypic and the phenotypic data for each trait for each individual using WinQTLCart software version 2.5. (Wang et al. 2012). The regions showing QTL peaks were scanned for associated SNP markers. Subsequently, the gene annotation database from the Sunup reference genome (http://www.plantgdb.org/XGDB/phplib/download.php?GDB=Cp) together with sequences within the NCBI BLAST server (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and the gene ontology database (http://www.geneontology.org) were utilised to determine the map location of associated sequences and putative candidate genes within the major QTL intervals.

4: Towards improvement of PRSV-P resistance in papaya

A resistant line BC4#113 which is a BC_4 of *V. pubescens* x *V. parviflora* was used as a resistance donor to cross to varieties of papaya '2.001', 'Sunrise Solo', 'RB1', 'RB2', 'RB4' and '1B'. Offspring of successful

crosses were embryo-rescued *in vitro*, multiplied and planted for evaluation of PRSV-susceptibility/resistance and fertility.

5: IDO component

The IDO was assigned with a 0.1FTE workload. This included traveling to and visiting all the major producers and field trials at various times throughout the project, to obtain a good understanding of all aspects of the papaya industry and research priorities. Organising and running an annual extension (Papaya Field Day) as well as other technical and scientific meetings and publishing in Hort Link.

Outputs

1. Advanced F4 breeding lines of red papaya (10 lines) and F3 yellow papaw (4 lines) were developed for two major growing regions (nine lines on the Tableland and five lines on the Coastal area). Seedlings of F4 red papaya and trees of F3 yellow papaw are to be evaluated in the next phase of the program.
2. A broad papaya germplasm was collected, databased and assessed for key phenotypic traits at three locations.
3. The Papaya Evaluation Handbook was developed and distributed for the standardized evaluation of productivity and fruit quality traits. The evaluation tools are being utilised by breeders to evaluate advanced breeding lines and as a communication tool with growers and marketers.
4. A comprehensive papaya genetic map was established with direct relevance to the Australian papaya industry. On this, the genome regions conditioning flesh sweetness, fruit firmness, flesh thickness, skin freckle and fruit size were identified and associated sequences determined. These markers need to be validated for further use to assist and accelerate the breeding program.
5. Seeds of F3 segregating populations for flesh sweetness were established for future genetic study.
6. Crosses between wild species containing PRSV-P resistance and papaya were successfully developed. Based on both DNA markers and morphological traits, these progenies were deemed to be true hybrids, fertile and contain the *prsv-1* (resistance) locus. These hybrids can be used as resistant donors to backcross to cultivated papaya. Collaboration with other leading organisations in Hawaii, India, Malaysia, Philippines, Taiwan and Thailand has been initiated for a more comprehensive study on PRSV-P resistance from wild species at the transcriptome level as well as sharing plant materials under MTA to successfully develop cultivated papaya to be resistant to the virus.
7. Activities targeted at industry development were organized to disseminate knowledge and research outcomes as follows:
    a. Three farm walks (field days):
        - On the 15th of April 2016, three farm visits in Innisfail followed by a seminar attended by 22 industry personnel.
        - On the 2nd of September 2016, two farm visits in Mareeba for Taiwanese researchers and a seminar workshop for growers and researchers attended by 20 industry personnel.
        - On the 22nd of September 2017, a farm visit in Mareeba followed by a seminar attended by 72 industry personnel.
    b. An article "R&D Snapshot on PP15000" was published in Hort Link 2018 Edition 1.
    c. Three articles were published in the mainstream media:

- Looking to grow sweeter papaya. Courier Mail, July the 28th, 2016.
- Can't keep our paws off papaya. Courier Mail, July the 29th, 2016.
- Pawpaw dreaming: try it, then love it. The Australian, July the 29th, 2016.

Outcomes

1. Currently, the market is restricted to eight varieties (five red papayas: RB1, RB2, RB4, Solo, Skybury; and three yellow papaws: 1B, H13, Costal Yellow). In recent testing, RB1 and Skybury were the preferred cultivars whereas 1B was the least preferred one (Tyler and Smyth, 2014). Skybury papaya is owned by Skybury Farmgate (no PBR) and appears to be highly susceptible to papaya sticky disease (SIAP meeting, 2018). RB1 is the standard red cultivar for the industry with a pleasant flavour and high yield, but poor skin quality (dull, prone to skin freckle).

   The advanced breeding lines developed in this project outperform these and the other industry standard varieties in many of the preferred agronomic and fruit quality traits, including setting fruit lower on the tree for ease of harvest, higher and consistent yields and sweet and pleasantly flavoured fruit. These lines can be used as new cultivars and/or new parental lines to develop new F1 hybrids. The development of novel varieties with superior fruit qualities will increase the domestic market and export potential of Queensland papaya.

2. The collection and preservation of papaya germplasm enables papaya breeders and researchers to access broader genetic materials than is currently available. The collections showed variation of traits in fruit size (300 g to 2.6 kg), peduncle length (short to very long), yield, sugar level (ºBrix level 8-14) and flesh texture (very soft to firm). Desirable characteristics such as high ºBrix level and long peduncle offer potential for future introgression into elite cultivars.

3. Standardised protocols (Papaya Evaluation Handbook) for assessing papaya productivity and fruit quality traits greatly improves the reliability of trait screening and selection by breeders leading to better continuity in the industry and improved communication between growers, marketers and researchers. It has been utilised throughout this breeding program and is used as a communication tool with growers (Amanda Arbuckle, Hayden Darvenisa, Gerard Kath, Mark McLaughlan, Eddie Mizzi and Joe Zappala), researchers (Dr Chat Kandana-udomkan, Mai Nantawan, Prof Rebecca Ford, Dr Paul Campbell, Dr Natalie Dillon, Yan Dizbalis and Lynton Vawdrey) and Hort Innovation personnel (Brad Mills and Vino Rajandran).

4. Knowledge related to papaya breeding, molecular breeding, bioinformatics and tissue culture of papaya has been transferred to papaya breeders, growers and researchers through scientific publications, industry communications and direct networking.

   Beneficial outcomes from the application of DNA markers to determine papaya seeding sex has impacted immediately on growers through cost reduction (reduced 50% of seed cost) and field labour, tree health and increase productivity. These outcomes also apply to other traits when molecular markers are developed for accurate and fast assisted selection. This project has identified potential QTL and genes for flesh sweetness, fruit weight, fruit length, fruit width skin freckle, flesh thickness and fruit firmness.

Micropropagation for papaya at a commercial scale is now possible. Knowledge has been transferred to Lecker Farming and Skybury Farmgate. Subsequently, more than 80% of their collective stock is now produced through the implementation of these tissue culture methods.

5. Knowledge on the genomic locations of major genetic components governing flesh sweetness, fruit firmness, flesh thickness, skin freckle, and fruit size has been produced. This is the foundation for the development of molecular trait selection tools for breeders, saving time and costs associated with traditional evaluation and selection. The identification and characterisation of sweetness-related candidate genes represents tools that may be immediately applied for selective sweetness breeding.

6. Development of papaya for Papaya Ringspot Virus (PRSV-P) resistance is a necessity to secure the Australian papaya industry. The first successful intergeneric hybridisation between *C. papaya* and *Vasconcellea* spp. using *V.parviflora* as a bridging species and an immune *V. pubescens* as a resistant donor was reported (Kanchana-udomkan et al. 2017). Based on both DNA markers and morphological traits, these progenies were deemed true hybrids, fertile and contained the *prsv-1* (resistance target locus). These F1 will be tested for resistance to PRSV-P and if so, will be used as a donor to transfer the resistance allele to cultivated papaya.

   The transfer of PSRV-P resistance to the cultigen is a major step forward in creating resistant superior varieties. The developed hybrid germplasm is ready for recurrent re-selective back-crossing to cultivated papaya.

7. This project has produced a trained Industry Liaison Officer (IDO) who is extremely familiar with all aspects of the Australian papaya industry and who has created multiple linkages with international papaya researchers. This has been the conduit to germplasm supply, and in communication with relevant R&D to growers and supply chain partners via meetings, workshops, media and direct communication.

   Also, the project has produced a PhD and Honours graduate that represent candidates for the new generation of papaya researchers and/or industry personnel. The project has also provided small research projects for five final year undergraduates, who contributed to the collection of vast amounts of phenotypic and/or genomic data.

These outcomes provide essential support to the Australian papaya industry through improved efficiency and sustainability via delivery of elite lines as well as novel tools for future selective breeding purposes. Once validated, the innovative and novel technologies for molecular-assisted selective breeding will significantly reduce the duration of selective breeding cycles and improve productivity. Knowledge in the areas of papaya breeding, molecular-assisted breeding, bioinformatics, tissue culture, disease response and consumer acceptability has been communicated to breeders, growers and researchers through scientific publications, industry communications and direct networking. Overall, the outcomes of the research program have strategically enhanced the development of the Australian Papaya Industry through operational excellence and strengthened the whole of industry networks.

Evaluation and Discussion

1. To improve eating quality especially in flesh flavour of yellow and red papaya and tailor-made varieties suitable to the main growing regions (Tableland and Coastal area).

The newly developed F3 selected lines performed significantly better than the current industry standards for yield, flavour, fruit flesh thickness and fruit skin blemishes (Table 4). Trees of selected F3 red papaya and yellow papaw lines are currently growing at multiple environmental trial sites. Seed of F4 of red papaya is ready to plant and that of yellow papaw will be harvested in 3-4 months.

Table 4: Mean and standard error of tree productivity and fruit quality traits of F2 yellow and F3 red breeding lines, compared to the industry standard varieties.

Line	Ranking	Sweetness (°Brix)	Fruit weight (kg)	Yield (kg)	Height to 1st fruit (cm)	Flesh thickness (cm)	Flesh ratio	Flesh colour (score)	Skin freckle (score)	Skin colour (score)	Skin quality (score)	Flesh firmness (score)	Fruit firmness (score)
1B	Std Variety	9.40 ± 0.49	1.27 ± 0.11	57.3 ± 5.2	107.8 ± 12	2.2 ± 0.1	0.55 ± 0.0	3.0 ± 0.0	1.2 ± 0.4	3.40 ± 0.8	1.78 ± 0.8	3.0 ± 0.0	3.0 ± 0.0
50-1 x 2-9	1	13.00 ± 1.00	1.48 ± 0.21	73.8 ± 10.0	69.5 ± 1	2.2 ± 0.2	0.51 ± 0.0	4.5 ± 0.5	3.5 ± 0.5	7.00 ± 0.0	3.00 ± 1.0	2.5 ± 0.5	2.5 ± 0.5
12-19	2	10.78 ± 0.63	1.32 ± 0.12	36.3 ± 2.6	72.8 ± 9	2.7 ± 0.4	0.59 ± 0.0	4.8 ± 0.6	2.0 ± 0.9	5.89 ± 1.0	3.44 ± 0.7	2.7 ± 0.5	2.5 ± 0.5
2-11 x 50	3	11.21 ± 1.36	1.30 ± 0.34	45.3 ± 4.9	71.5 ± 16	2.3 ± 0.4	0.55 ± 0.0	3.6 ± 0.7	1.6 ± 0.7	5.86 ± 1.0	3.29 ± 0.7	1.6 ± 0.0	2.0 ± 0.0
12-7	4	11.56 ± 0.73	1.25 ± 0.22	43.0 ± 4.0	74.3 ± 3	2.3 ± 0.2	0.60 ± 0.0	3.5 ± 0.5	2.5 ± 0.5	6.25 ± 1.3	2.88 ± 0.3	1.5 ± 0.0	2.0 ± 0.0
RB1	Std Variety	11.41 ± 1.91	0.92 ± 1.22	14.9 ± 8.0	81.7 ± 15	2.3 ± 0.0	0.64 ± 0.0	7.1 ± 0.5	2.6 ± 1.0	6.46 ± 0.4	2.35 ± 0.4	1.5 ± 0.7	2.3 ± 0.3
T-1	1	12.80 ± 1.20	1.06 ± 0.22	36.1 ± 0.0	90.0 ± 0	2.5 ± 0.2	0.65 ± 0.1	6.0 ± 0.3	1.0 ± 0.9	6.00 ± 0.7	3.00 ± 0.2	3.0 ± 0.0	3.0 ± 0.0
T-2	2	12.80 ± 0.55	0.90 ± 0.35	27.1 ± 0.0	62.0 ± 0	2.6 ± 0.1	0.74 ± 0.1	7.0 ± 0.5	1.0 ± 0.8	7.00 ± 0.2	2.00 ± 0.5	2.0 ± 0.0	3.0 ± 0.0
T-3	3	12.00 ± 0.65	0.94 ± 0.26	20.8 ± 0.0	45.0 ± 0	2.5 ± 0.2	0.59 ± 0.1	6.0 ± 0.5	2.0 ± 0.5	7.00 ± 0.9	4.00 ± 0.3	2.0 ± 0.1	3.0 ± 0.0
I-1	1	14.20 ± 1.20	1.17 ± 0.26	21.1 ± 0.0	72.0 ± 0	2.7 ± 0.3	0.68 ± 0.2	8.0 ± 0.3	2.0 ± 0.7	6.00 ± 0.7	2.00 ± 0.2	3.0 ± 0.2	3.0 ± 0.0
I-2	2	13.70 ± 0.80	0.94 ± 0.33	16.0 ± 0.0	38.0 ± 0	2.2 ± 0.3	0.64 ± 0.1	7.0 ± 0.4	2.0 ± 0.8	6.00 ± 1.1	2.00 ± 0.7	2.0 ± 0.0	3.0 ± 0.0
I-3	3	13.10 ± 0.78	1.21 ± 1.22	12.1 ± 0.0	92.0 ± 0	3.0 ± 0.5	0.71 ± 0.4	7.0 ± 0.5	2.0 ± 0.7	7.00 ± 0.5	4.00 ± 0.3	3.0 ± 0.1	2.0 ± 0.5

2. To be able to replicate the new elite lines by seed production (priority) and micropropagation.

The selected F3 breeding lines which were evaluated within this project were genetically unstable. Hence, they were self-pollinated to stabilise their genetics (94% stable) and F4 generations were established. Two lines of yellow and four lines of red papaya were successfully developed and maintained in tissue culture. Seed was produced from all lines.

3. To understand the performance of exotic genetics in different growing areas - leading to selection of characteristics to transfer to Australian genetics and generate a database of papaya genetics for future reference.

The procedure to import seed from overseas has taken longer than expected as this has to include a negotiated Material Transfer Agreement from all parties. Many of the seed received in this project had very poor germination and some failed to germinate. This is likely due to improper storage prior to transfer from their home facility as well as the potential for rapidly deterioration of papaya seed after harvest (Sangakkara 1995). A database of the collected germplasm was generated that, wherever possible, included phenotypic data of fruit quality and productivity (Appendix 4) along with a DNA profile generated using 48 SSR primers (Appendix 5, Figure 1). Within 18 genotypes, 90 alleles were detected. More than 80% of these alleles were present in Skybury, RB4 and Linda Solo, while only 40% were detected in RB1

and RB2, demonstrating their genetic differentiation. The DNA profiles constructed using SSR markers differentiated among all of the genotypes included. In addition, DNA of commercial Australian cultivars (RB1, RB2, RB, 1B, Solo, ER62-yellow parental line) and a Thai cultivar (Holland aka 'Sekaki') were genotyped by SNP detection via GBS to create robust molecular fingerprints/profiles to be used in PBR protection (Appendix 6). Interestingly, the genotypes grouped based on colour and origin. Genotypes varied in their growth and fruit quality traits across the locations in which they were assessed. In general, trees in Innisfail grew taller, bigger and had longer peduncles than those in Mareeba (Table 5).

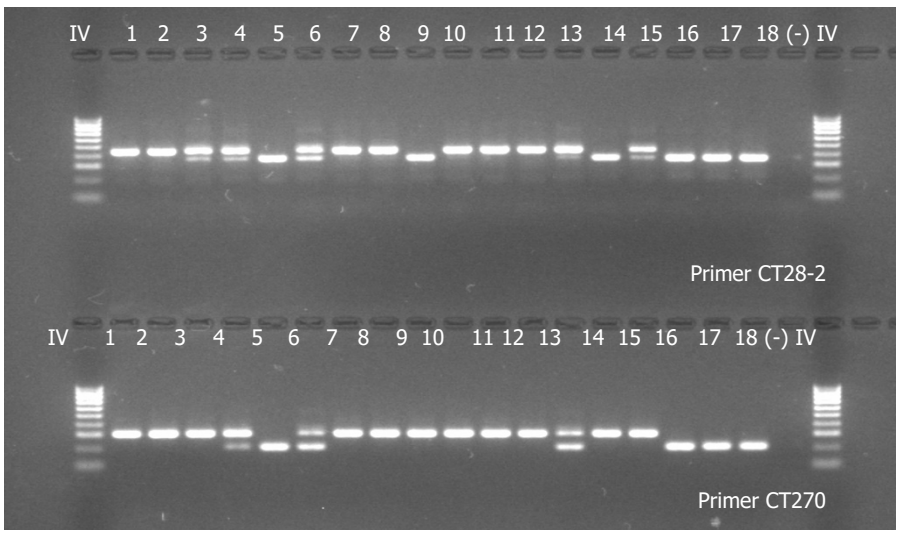

Figure 1: Gel photo showing an example of 18 (1-18) genotype profiles of *C. papaya* produced using two SSR primers (CT28-2 and CT270)

Table 5: The mean values of four productivity traits measured at two locations (JZ = Joe Zappala, Innisfail; GK = Lecker Farming, Mareeba), five months after field planting.

Line	Source	Height to 1st fruit (cm)			Trunk size (cm)			#Side shoot			Peduncle length score		
		JZ[1/]	GK[1/]	P-Value[2/]	JZ[1/]	GK[1/]	P-Value[2/]	JZ[1/]	GK[1/]	P-Value[2/]	JZ[1/]	GK[1/]	P-Value[2/]
GR-01	**Australia**	101.0 abc	67.4 bc	0.002	40.6 abc	36.6 a	0.008	0.0 c	0.0 b	1.000	2.4 ab	1.4	0.020
GR-02	**Australia**	82.6 abc	77.8 ab	0.493	43.3 ab	35.6 a	0.018	0.0 c	0.0 b	1.000	3.0 a	1.6	0.005
GR-03	**Australia**	107.3 abc	71.4 bc	0.024	45.7 a	36.4 a	0.021	2.6 a	0.0 b	0.071	2.0 ab	1.4	0.116
GR-04	**Hawaii**	123.3 a	90.2 a	0.011	36.0bcd	31.2 ab	0.206	0.3 b	2.4 a	0.189	3.0 a	1.0	0.000
GR-07	**Thailand**	84.3 abc	65.2 bc	0.024	26.7 e	30.0 ab	0.153	0.0 c	0.0 b	1.000	3.0 a	1.2	0.001
GR-09	**Taiwan**	108.0 abc	68.8 bc	0.003	35.7 bcd	27.6 b	0.030	2.0 a	0.0 b	0.003	1.5 b	1.2	0.411
GR-10	**Taiwan**	111.3 ab	68.2 bc	0.001	36.0 bcd	31.8 ab	0.011	0.3 b	0.0 b	0.220	3.0 a	1.2	0.001
GR-21	**Australia**	87.6 abc	93.2 a	0.703	40.7 abc	29.0 ab	0.010	0.0 c	0.0 b	1.000	2.3 ab	1.0	0.005
GR-23	**Australia**	72.5 bc	68.0 bc	0.308	31.2 de	25.4 b	0.005	0.5 b	0.0 b	0.292	2.2 ab	1.6	0.109
GY-24	**Australia**	98.5 abc	81.2 ab	0.042	40.5 abc	27.2 b	0.007	0.0 c	0.0 b	1.000	3.0 a	1.0	0.000
GY-26	**Australia**	77.6 bc	92.8 a	0.155	34.7 cde	31.4 ab	0.438	0.0 c	0.0 b	1.000	2.6 a	1.4	0.021
GY-29	**Australia**	67.7 c	58.6 c	0.107	33.0 cde	35.6 a	0.029	0.0 c	0.0 b	1.000	3.0 a	1.0	0.000
P-Value[3/]		0.000	0.000	0.000	0.000	0.000	0.000	0.018	0.000	0.126	0.000	0.281	0.000

[1/] Means followed by the same letter in each column are not significantly different

[2/] P-Values in each column were obtained from T-test of data between the two farms

[3/] P-Values in this row were obtained from ANOVA of data within the same farm comparing among germplasm lines

4. To establish breeding and segregating populations for traits identified by the industry and consumers, such as skin and flesh quality traits.

F2 segregating populations were developed to study the genetics conditioning fruit quality traits including flesh flavour, flesh texture, flesh thickness, flesh sweetness, flesh consistency, skin quality, skin colour, and peduncle length (Table 6). Five F3 segregating populations for flesh sweetness were established. These populations can be used to study genetic inheritance and develop associated molecular markers.

Table 6: The lines developed for each trait of interest within the project.

Number of Lines	Description	Traits of study
2	F2 (RB2 X Sunrise Solo)	Flavour
3	F2 (RB2 x 1B)	Flavour
3	F2 (Linda solo x (RB1x 18_45)	Skin quality
2	F2 (2.54-14 self x (JC2xVietnam Red))	Skin colour
2	F2((JC2xVietnam Red) x 2.54-14 self)	Skin colour
2	F2 ((7_82 x 1B) x 2.54-12 self)	Flesh colour, texture, thickness, sweetness
1	F2 ((7_82 x 1B) x (24-87x 1B)	Flesh consistency
3	F2 (Red flesh unknown 2 x (RB2 x 18-45))	Flesh color
4	F2 (RB2 x Linda Solo)	Flesh texture, thickness
4	F2 (Linda Solo x RB2)	Flesh texture, thickness
3	F2 (RB2 X Sunrise Solo)	Sweetness
1	F2 ((25-5(x RB1) x (RB2 x 18-45))	Peduncle length
5	F3 (RB2 x Sunrise Solo)	Sweetness

5. To develop bioassays for robust, reliable and practical evaluation of fruit quality traits.

Standard evaluation methods for papaya fruit quality and tree productivity traits were developed and published (Nantawan et al. 2017). Harvesting indices were also optimised by investigating key fruit quality traits within a time course of fruit maturity stages (Figure 2) and post-harvest. Fruit firmness, flesh firmness, skin colour, consistency of flesh colour and flesh sweetness were distributed along with harvest maturity variables (Figure 3). Considering the correlation between these quality traits and harvest maturity, maturity stages 1 and 2 were associated with fruit firmness and flesh firmness. On the other hand, maturity stages 3 to 5 were associated with sweetness, skin colour and consistency of flesh colour. As a result, harvesting at Maturity Stage 3 was shown to robustly lead to a fully ripened fruit with the highest quality characteristics at 6 days after harvest (Nantawan 2018). The standard evaluation methods and harvesting indices were subsequently applied to all evaluations throughout the project.

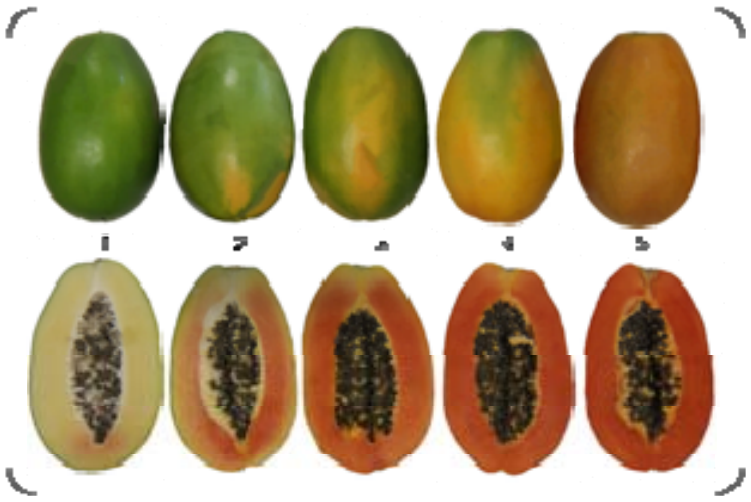

Figure 2: Five fruit maturity stages

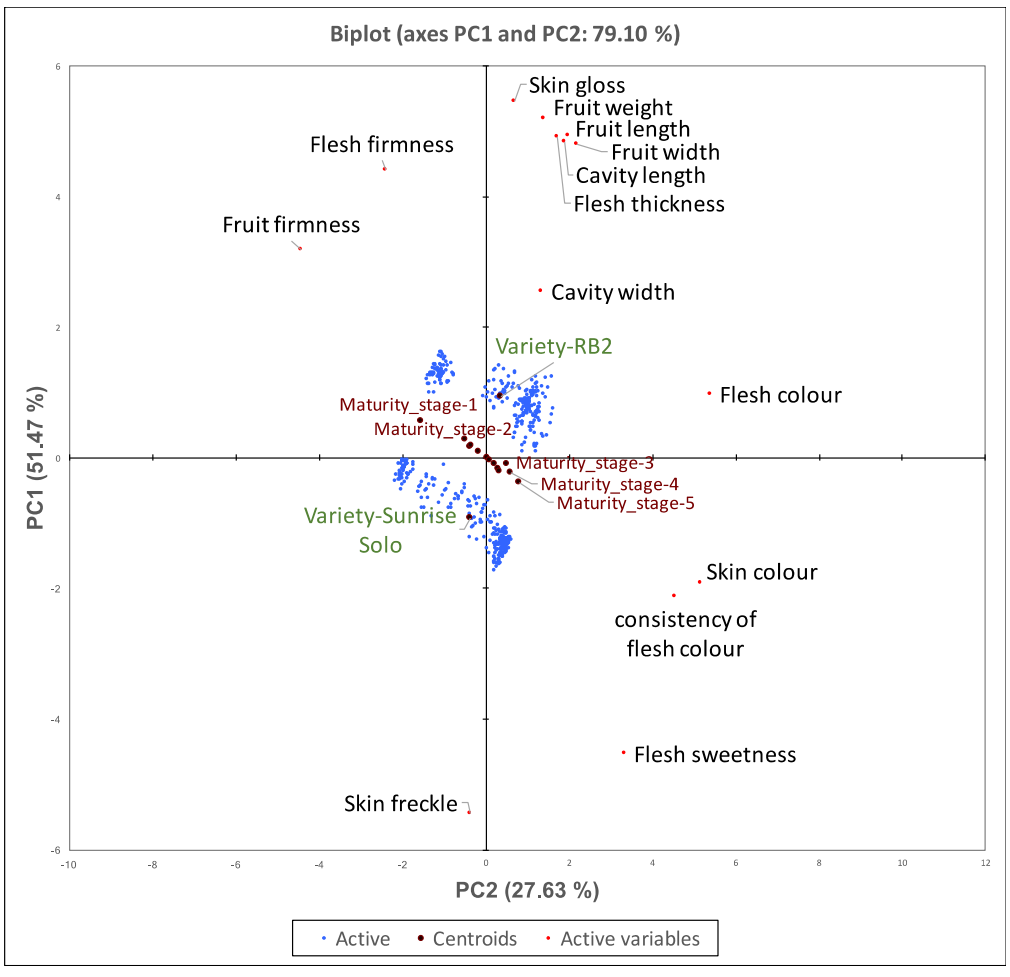

Figure 3: Correlation among fruit quality traits based on principle component analysis of data from multiple sites across two harvest years.

6. To study genetic inheritances of selected quality traits from the developing populations.

The values of all traits assessed were significantly different between the parents used in this project (Figure 4). The F1 progeny had mean values that fell between those of the parents and the F2 progeny had a range of values within the range of the two parents.

Moderate to high heritability was detected for most of the fruit quality components except consistency of flesh colour, skin gloss, skin colour and fruit cavity. This therefore are proposed to be more influenced by environmental factors. High heritability (over 70%) was detected for flesh colour, fruit firmness, flesh firmness and flesh sweetness, indicating that these traits can be more readily exploited through genetic approaches for improving fruit quality in papaya.

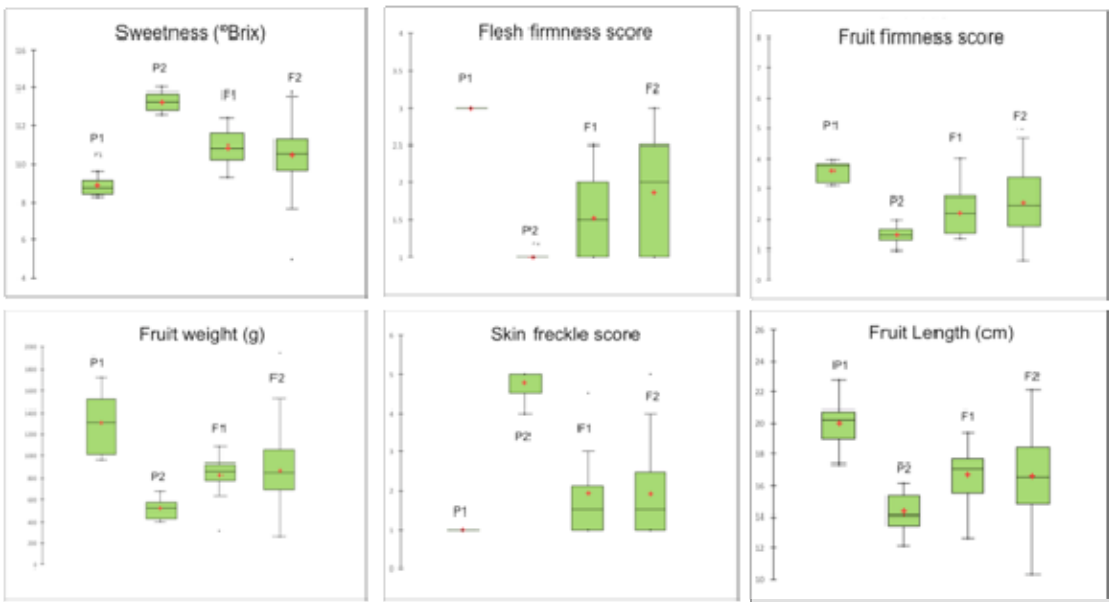

Figure 4. Boxplot of sweetness, flesh firmness, fruit firmness, fruit weight, skin freckle and fruit length of parental varieties (P1 = RB2 and P2 = Sunrise Solo), their F1 and F2 population. The bars represent standard errors, the boxes represent distribution of data and the + represents mean value.

7. To identify and develop DNA markers using genome sequencing techniques - to develop new DNA markers as well as using microsatellite markers from existing databases to create a dense framework map.

A genetic map-based QTL analysis was performed to identify important genomic regions linked to flesh sweetness and other key fruit traits. The genetic map developed from this project was based on GBS-generated SNP polymorphisms. The final linkage map was constructed from 219 highly robust SNP markers and comprised 11 linkage groups covering a total distance of 509 cM. A total of 21 QTL were identified for seven key fruit quality traits including flesh sweetness, fruit weight, fruit length, fruit width skin freckle, flesh thickness and fruit firmness characters (Appendix 3). These QTL independently explained a significant amount of the phenotypic variation (6.5 – 35.8%) of each trait (Table 7). Nine major QTL (explaining more than 10% phenotypic variance) and seven stable QTL detected across two harvest years related to flesh sweetness, fruit weight, length, width and firmness were identified, Candidate genes for these major QTL were proposed (Table 8). Those associated with flesh sweetness were related to nucleotidyltransferase activity, cell division, DNA replication, cellular response to DNA damage stimulus, vesicle-mediated

transport, intracellular protein transport and protein transmembrane transporter activity. Once validated, the markers/genes that are significantly associated to the stable major QTL present a large opportunity for the application of marker-assisted trait breeding.

Table 7: Summary of QTL detected for fruit quality traits

Trait	Year in which QTL was detected	Number of QTL detected	% of phenotypic variance explained
Flesh sweetness	SWE_2016	1	11.6
	SWE_2017	2	13.9
Fruit weight	WEI_2016	3	31.5
	WEI_2017	3	30.4
Fruit length	LEN_2016	2	29.5
	LEN_2017	2	25.6
Fruit width	WID_2016	1	19.5
	WID_2017	4	29.1
Skin freckle	FRE_2016	2	15.95
	FRE_2017	3	16.95
Fruit firmness	FIR_2016	5	34.2
	FIR_2017	4	35.8
Flesh thickness	THI_2016	2	12.7
	THI_2017	1	6.5

Table 8: Associated SNPs and putative candidate genes identified for fruit quality traits.

Trait	Marker	%Var 1/	Candidate genes	Gene functional annotation	Candidate genes GO terms	Corresponding gene function
Flesh sweetness	sCT_80_454708	11.6	CP00080G00550	Non-canonical poly(A) RNA polymerase PAPD5	GO:0016779, GO:0051301; GO:0071044	Nucleotidyltransferase activity, cell division, Histone mRNA catabolic process
			CP00080G00540	KIN17-like protein	GO:0006260; GO:0006974; GO:0048638	DNA replication; cellular response to DNA damage stimulus, regulation of developmental growth
	sCT_12_1083429	10.6	CP00012G01230	Transmembrane emp24 domain-containing protein P24-beta2	GO:0016192, GO:0006886, GO:0008320	Vesicle-mediated transport; intracellular protein transport, protein transmembrane transporter activity
Skin freckle	sCT_33_864201	8.5	CP00033G00820	Ultraviolet-B receptor UVR8	GO:0009881, GO:0006281	Photoreceptor activity, DNA repair
	sCT_16_2143939	3.23	CP00016G01460	Putative disease resistance protein RGA1	GO:0006952, GO:0009723, GO:0009737	Defence response, response to ethylene, abscisic acid

Table 8: (Cont.)

Trait	Marker	%Var [1/]	Candidate genes	Gene functional annotation	Candidate genes GO terms	Corresponding gene function
Fruit firmness	sCT_6_237757	6.1	CP00006G00370	UPF0553 protein-like	GO:0005575; GO:0006400	Cellular component; tRNA modification
	sCT_7_2121986	4.3	CP00007G02090	DNA-directed RNA polymerase III subunit 1	GO:0001056	RNA polymerase III activity
	sCT_48_1243956	4.4	CP00048G02250	MYB-like protein X	GO:0003677; GO:0001135	DNA binding; transcription factor activity, RNA polymerase II transcription factor recruiting
			CP00048G02260	Pectin acetyl esterase 12-like	GO:0045490	Pectin catabolic process
Fruit weight	sCT_20_560108	14.2	CP00020G00540	Protein trichome birefringence-like 12	GO:0071554	Cell wall organization or biogenesis
	sCT_6_2754743	12.7	CP00006G03490	Serine/threonine-protein kinase/endoribonuclease IRE1	GO:0004674	Protein serine/threonine kinase activity
			CP00006G03500	Glutamate receptor 3	GO:0008066	Glutamate receptor activity
	sCT_20_544712	13.9	CP00020G00510	Fatty acid amide hydrolase-like	GO:0017064; GO:0009505; GO:0004128	Fatty acid amide hydrolase activity; plant-type cell wall; cytochrome-b5 reductase activity, acting on NAD(P)H
	sCT_6_2392635	11.3	CP00006G03260	IST1-like protein	GO:0015031	Protein transport
Fruit length	sCT_50_1447788	12.3	CP00050G01530	Exopolygalacturonase-like	GO:0016021; GO:0005975	Integral component of membrane; carbohydrate metabolic process
	sCT_42_954411	19.8	CP00042G00970	NAC domain-containing protein 41-like	GO:0005975	Carbohydrate metabolic process
			CP00042G00960	*Carica papaya* chromosome Y sequence		
			CP00042G00980	Prolyl 4-hydroxylase 9	GO:0019538	Protein metabolic process
Fruit width	sCT_6_2331252	11.3	CP00006G03140	Bifunctional nuclease 2	GO:0032296	Double-stranded RNA-specific ribonuclease activity
Flesh thickness	sCT_6_1666511	8.4	CP00006G02040	Cytochrome P450 84A1-like	GO:0009809, GO:0016020	Lignin biosynthetic process, membrane,
			CP00006G02050	Protein FMP32, mitochondrial	GO:0033617	Mitochondrial respiratory chain complex IV assembly
	sCT_114_766126	4.3	CP00114G00550	Ethylene-responsive transcription factor RAP2-7-like	GO:0009873	Ethylene-activated signalling pathway

[1/] % Explained phenotypic variance

8. To develop papaya with resistance to PRSV-P by introgression of resistance genes from *Vasconcellea* species.

One of the heterozygous lines, BC4#113, was selected as a resistance donor to cross with *C. papaya* line 2.001. Embryos were rescued, initiated and multiplied *in vitro*. The cross was successfully micropropagated

and produced 18 F1 lines [*C. papaya* x (*V. parviflora* backcrossed to *V. parviflora* x *V. pubescens*)]. The first successful intergeneric hybridisation between *C. papaya* and *Vasconcellea* spp. using *V. parviflora* as a bridging species was reported (Kanchana-udomkan et al. 2017). Based on both DNA markers and morphological traits (Figure 5), these progenies were deemed true hybrids, fertile and contained the *prsv-1* (resistance target locus). These F1 will be tested for resistance to PRSV-P and if so, will be used as a donor to transfer the resistance allele to cultivated papaya.

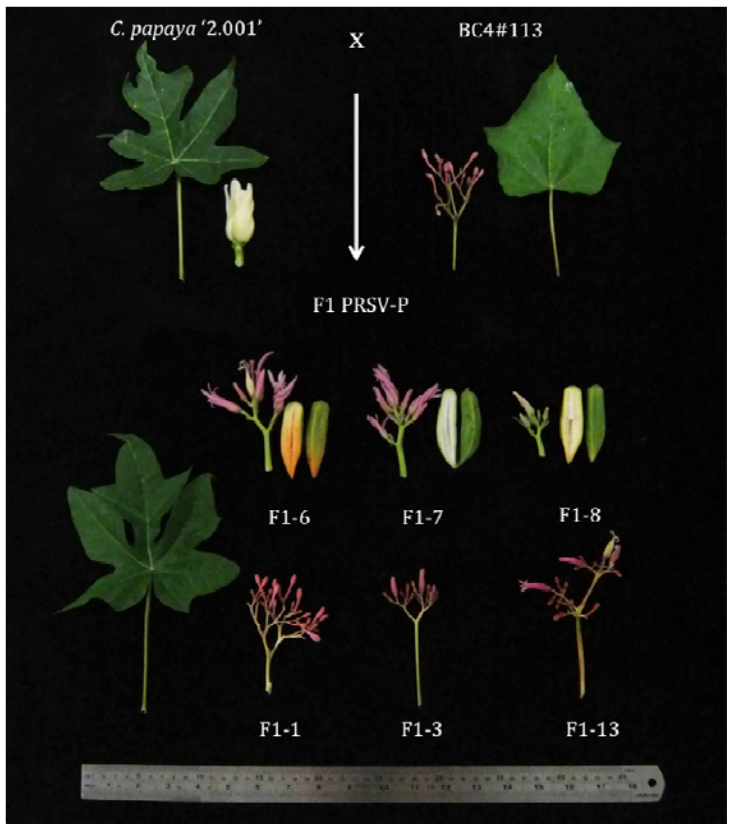

Figure 5: Leaves and flowers of parental lines (*Carica papaya* line 2.001 and BC4#113) and their F1 progenies. F1 lines 6, 7 and 8 show female flowers and fruits, and lines 1, 3 and 13 show male flowers.

Recommendations

- Speed breeding technology should be investigated to accelerate the breeding cycle and shorten the time to cultivar release.
- The new stable genetic materials should be progressed to cultivars and be used in the breeding program as new parental lines.
- The broader germplasm collection should be planted and maintained on a research property for further phenotypic assessment.
- Sweetness-related genes identified from Sunrise Solo should be assessed as suitable sweetness markers and possible stable transfer to other and newly developed elite advanced breeding lines.
- The flavour and taste profiles of the advanced breeding lines should be profiled in depth and in synergy with consumer-driven assessment to ensure the best preferences are selected.
- SNP markers associated to other fruit quality trait QTLs should be validated and developed into transferable markers to assist in the selection process to other and newly developed elite advanced breeding lines.
- All advanced breeding germplasm should be assessed for disease reaction to major papaya diseases to inform expectation for their future Integrated Disease Management.

Scientific Refereed Publications

Journal articles

Kanchana-udomkan, C., Drew, R.A., Ford, R., in press. Molecular marker-assisted papaya sex determination for improved grower efficiency. *Acta Horticulturae*

Nantawan, U., Kanchana-udomkan, C., Drew, Ford R., in press. Development of polymorphic simple sequence repeat (SSR) markers from genome re-sequencing of *Carica papaya* L. `Sunrise Solo´ and `RB2´ for marker- assisted breeding. *Acta Horticulturae*

Kanchana-udomkan, C., Nantawan, U., Drew, R., Ford, R., in press. Progress in Introgression of Papaya Ringspot Virus Resistance from *Vasconcellea pubescens* to *Carica papaya* L. *Acta Horticulturae*

Nantawan, U., Kanchana-udomkan, C., Drew, Ford R., in press. Identification of genes related to sugar accumulation in *Carica papaya* L.: differentially expression and candidate markers development. *Acta Horticulturae*

Kanchana-udomkan, C., Drew, R.A., Ford, R., 2016. A kinase gene potentially implicated in resistance to Papaya ringspot virus in *Vasconcellea quercifolia*. *Acta Horticulturae* **1111**, 41-48.

Nantawan, U., Kanchana-udomkan, C., Drew, R., 2016. Progress in Marker Assisted Breeding of Papaya in Australia. *Acta Horticulturae*. **1111**, 79-86.

PhD thesis

Nantawan, U., 2018. Towards marker-assisted breeding of important physiological and fruit quality traits in papaya (*Carica papaya* L.). PhD thesis, Griffith University, Nathan, 193 pages

Intellectual Property/Commercialisation

Pre-Existing IP:

Griffith University have an MTA with DAFQ for 6 lines (as noted previously in PP10005) and recognize that Papaya SSR molecular markers for Virus resistance are available in the public domain – all sequences are published and no IP claims have been made by the researchers/developers in the USA. Research Leader's understanding, there is no issues in using them including for the delivery of commercial lines to Australian growers.

Intellectual Property, Commercialisation and confidentiality from this project:

The germplasm that has been collected includes 3 lines from Thailand (Khaeg dum, Khaeg nuen and Sekaki) and 2 lines from Taiwan (Tainung #2 and First lady) was supplied without an MTA. Potential for claims on IP ownership requires further exploration. The commercial lines sourced from Australia are originally licensed by Papaya Seed Australia. We will explore further the current licensing agreements/ownership for all lines in the germplasm collection and arrange MTA and license to use where relevant.

The protection of new IP generated through the production of the advanced breeding lines and future protection through to commercialization has been discussed with Griffith Enterprise and Horticulture Innovation. The plan forward is to generate Plant Breeders' Rights (PBR) for these materials inclusive of the molecular profiles generated.

Genomic data of whole genome sequences of *Carica papaya* '1B', 'RB2' and 'Sunrise Solo', *Vasconcellea pubescens*, *V. quercifolia* and *V. parviflora* have not been placed in the public domain yet. Discussion will be sought with HI to do so. Some of the whole genome sequences were provided to the National Taiwan University under an agreement (Appendix 7) for the purposes of collaboratively seeking and tracking functionality of the PSRV-P resistance alleles.

References

Horticulture Innovation Australia (HIA), 2017. Papaya Strategic Investment Plan 2017 – 2021

Kanchana-udomkan, C., Nantawan, U., Drew, R., Ford, R., 2017. Progress in Introgression of Papaya Ringspot Virus Resistance from *Vasconcellea pubescens* to *Carica papaya* L. *The 4th International Symposium on Molecular Markers in Horticulture*: 7-10 March 2017, Napier, New Zealand

Ming, R., Hou, S., Feng, Y., et al., 2008. The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature 452:991–996. doi: 10.1038/nature06856

Nantawan, U., 2018. Towards marker-assisted breeding of important physiological and fruit quality traits in papaya (*Carica papaya* L.). *PhD thesis*, Griffith University, Nathan, 193 pages

Nantawan, U., Kanchana-udomkan, C., Ford, R., 2017. Papaya Evaluation Handbook. 24 pages. https://horticulture.com.au/wp-content/uploads/2017/09/Papaya-Evaluation-Handbook.pdf

Sangakkara, U.R., 1995. Influence of seed ripeness, sarcotesta, drying and storage on germinability of papaya (*Carica papaya* L.) seed. *Pertanika Journal of Tropical and Agricultural Science*. **18** (3), 193-199

Tyler, P., Smyth, H., 2014. Identifying consumer preferences for papaya and papaw with view to increase value and consumption. Final Report – Executive summary, p4

Wang, S., Basten, C.J., Zeng, Z.B. 2012. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, <http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>

Acknowledgements

We recognise the valuable contributions made by the following growers who provided land, time, resources and experience to aid in the trial plantings on their properties:

1. Gerard Kath
2. Ian McLaughlan
3. Eddie Mizzi
4. Michael Oldano
5. Jeff and Sirin Slucher
6. Joe Zappala

We thank the members of the Advisory Group who met regularly with the project research team and contributed valuable inputs towards breeding objectives and evaluation of trial plantings. We thank the following members of this group:

1. Hayden Darvenisa
2. Gerard Kath
3. Mark McLaughlan
4. Joe Zappala

Appendices

1. Kanchana-udomkan, C., Nantawan, U., Drew, R., Ford, R., in press. Progress in Introgression of Papaya Ringspot Virus Resistance from *Vasconcellea pubescens* to *Carica papaya* L. Acta Horticulturae
2. List of SSR primers
3. Database of papaya germplasm
4. DNA profile of papaya germplasm (full size figure attached)
5. DNA Finger-printing of Commercial Australian Papaya Cultivars by Genome-by-Sequencing Analysis
6. Agreement for the sharing of *Vasconcellea* spp. and *Carica papaya* sequences owned by Griffith University with National Taiwan University

Progress in introgression of papaya ringspot virus resistance from *Vasconcellea pubescens* to *Carica papaya*

C. Kanchana-udomkan[1,2], U. Nantawan[1,2], R. Drew[1] and R. Ford[1,2]

[1]School of Natural Sciences, Griffith University, Nathan, QLD; [2]Environmental Futures Research Institute, Griffith University, Gold Coast, QLD

Abstract

An attempt to develop *Carica papaya* L. resistant to papaya ringspot virus type P (PRSV-P) has been established via an intergeneric cross using *Vasconcellea pubescens* as the resistance donor. *V. pubescens* has consistently been reported to be resistant to PRSV-P unlike the resistance in other species (d'Eeckenbrugge et al., 2013). The immunity to PRSV-P is also controlled by a single dominant gene (Dillon et al., 2006), which should be easly to transfer across crops. Due to the genetic distance between the two genera and genetic incompatibility, *V. parviflora* was introduced as a bridging species to transfer PRSV-P resistance gene(s) to papaya (*C. papaya*). Backcrossing has been achieved to BC4 in order to maximise the *V. parviflora* genetic background to improve fertility. The PRSV-P resistance genotype of the BC4 was determined with previously published sequences linked to the *prsv-1* locus (Dillon et al., 2006). One of the heterozygous lines, BC4#113, was selected to use as a resistance donor to cross with *C. papaya* line 2.001. Embryos were rescued, initiated and multiplied *in vitro*. The cross was successfully micropropagated and produced 18 F1 lines [*C. papaya* x (*V. parviflora* backcrossed to *V. parviflora* x *V. pubescens*)]. These F1s were genotyped at the *prsv-1* locus as previously mentioned. They were subsequently acclimatised and planted in a field in north Queensland. True hybridity was revealed through morphological characterisation, including pink flowers, which is a dominant trait from *V. parviflora*. The hybrid F1s carrying the PRSV-P locus were then successfully backcrossed to *C. papaya* to generate BC1 generations. They were also sib-crossed with the F1 to generate F1:F2 generations. These lines will be planted in a quarantine laboratory at Griffith University, Nathan for virus resistance evaluation. Fertile plants that carry the virus resistance alleles will be used as parental lines in future breeding strategies.

Keywords: PRSV-P resistance, interspecific hybridisation, intergeneric hybridisation, papaya breeding, embryo rescue

INTRODUCTION

Papaya ringspot virus disease, caused by papaya ringspot virus type P (PRSV-P), is a major issue in papaya cultivation worldwide. For several decades, the disease has been known to cause economic damage affecting papaya plantations across growing regions (Purcifull et al., 1984), including Australia (Thomas and Dodman, 1993). It results in rapid decline of plant growth and fruit development. The infected plants often yield unmarketable fruit and die (Gonsalves, 1998). To date, a number of methods to combat the disease have been attempted, such as cultural practices, cross-protection and crop improvement. However, these methods do not suppress PRSV-P invasion and do not control the disease in the long-term. By far the most effective method for long-term prevention is the use of PRSV-P resistant cultivars (Gonsalves, 1998). Genetically modified resistant papaya cultivars have been used in some countries (Bau et al., 2003, 2004; Fitch et al.. 1993; Kung et al., 2009; Tripathi et al., 2008). However, issues related to public perception and the countries' regulatory procedures have limited their use.

A conventional breeding approach is preferred and progress in improvement of papaya resistance to PRSV-P by this method has to date been promising. The natural PRSV-P

resistance found in wild relatives of papaya such as the genus *Vasconcellea*, is realistically heritable (Manshardt and Wenslaff, 1989; Magdalita et al., 1997; Drew et al., 1998a). In Australia, the papaya breeding program aims to utilise the immunity from *V. pubescens* (Drew et al., 1998a; Sharma and Tripathi, 2013). However, attempts to move the resistance to papaya was obstructed by intergeneric hybridization barriers, which resulted in low numbers of progenies and all progenies being infertile (Drew et al., 1998). To overcome crossing-incompatibility, another species, *V. parviflora*, was introduced as a bridging species to transfer PRSV-P resistance from *V. pubescens* into *C papaya*. Interspecific hybridisation between *V. pubescens* and *V. parviflora* produced fertile homozygous resistance (RR) in the F2 generation (O'Brien and Drew, 2010). However, when crossing to papaya was attempted, the hybrids failed to produce fertile flowers, likely because the F2 generation was largely *V. pubescens* and maintained the fertility barriers (Drew et al., 1998). As a result, a backcrossing program to increase the genetic content of *V. parviflora* within the hybrids was conducted. The plants from each backcross (BC) generation were selected for PRSV-P resistance using the Opk4_1r SCAR marker and *Psi*Ik4 CAPS markers, which are co-located within the *prsv-1* locus (Dillon et al., 2006). Next, the interspecific hybrids containing the *prsv-1* locus were backcrossed to papaya from which the resistance line #113 (BC4#113) was selected as it produced viable pollen (Drew et al., 2005).

This study aims to examine the intergeneric hybridisation between *C. papaya* and BC4#113, assess the F1 population for fertility and initiate breeding lines for further crossing in the PRSV-P resistance breeding program.

MATERIALS AND METHODS

Plant materials

1. Intergeneric and interspecific hybridisation

BC4 populations were produced from the original cross of *V. pubescens* and *V. parviflora* (Figure 1) from previous papaya breeding research at Griffith University, Nathan, Australia (O'Brien and Drew, 2009; Razean Haireen, 2013). One plant of the BC4 population, line #113 (BC4#113), was selected to be the resistant donor and the male parent in a cross with *C. papaya* line 2.001. The 2.001 genotype previously showed high cross-compatibility with *Vasconcellea* species and a high yield of embryos (Drew et al., 2005). Fruits were harvested at an immature stage (90-100 days after pollination) to perform embryo rescue.

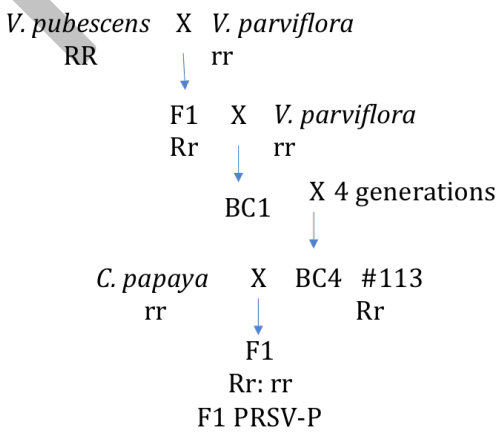

Figure 1. Diagram of bridge-breeding program for papaya ringspot virus-P resistance by introgression of resistance genes from *Vasconcellea pubescens* to *V. parviflora*, then outcrossed to *Carica papaya*. The alleles (R, r) presented in the diagram were identified by *Psi*Ik4 CAPS marker (Dillon et al., 2006).

2. Micropropagation

Seed from immature fruit were harvested and embryo rescue was performed following the protocol described by Drew et al. (1998) and Razean Haireen (2013), and the individual embryos were inoculated onto induction medium (Drew and Smith, 1986). The established embryonic cultures were micropropagated using protocols described by Drew (1992) to increase the quantity of the plant materials for further experimentation. Once adequate numbers of plants were generated, the plants were acclimatised following the procedures outlined by Drew (1988) and grown to maturity.

Molecular analysis in F1 generation

1. Hybridity assessment

Genomic DNA extraction

Fresh young leaves of *in vitro* plantlets were collected for DNA extraction using a modified cetylmethylammonium bromide method (Dellaporta et al., 1983). DNA quality and quantity was assessed by spectrophotometer (NanoDrop 1000c, Thermo Scientific) and diluted to 25 ng/µL for PCR analysis.

PCR assay

DNA of the parents (line 2.001 and BC4#113), the original parents of line BC4#113 (*V. pubescens* and *V. parviflora*) and 18 lines of F1 progeny was assessed. The simple sequence repeat (SSR) primers were selected on their ability to amplify and differentiate between the parental DNAs. They were located close to the *prsv-1* locus on linkage group 5 (LG5), the region of interest for PRSV-P resistance (Kanchana-udomkan et al., 2015).

PCR reactions were performed in 15 µL volumes, containing 2 µL of genomic DNA (25 ng/µL), 0.3 unit MyTaq Red DNA Polymerase (Bioline), 1X MyTaq Red Reaction Buffer (1 mM dNTPs, 3mM $MgCl_2$, stabilizers and enhancers; Bioline) and 0.3 µM each of forward and reverse primers (Table 1, primers 1 to 3). PCR cycles were carried out in a thermal cycler (MyCycler, BioRad) using a two-step PCR. The program was set at an initial denaturation of 95°C for 10 mins, then the reaction was repeated for the first annealing of 48°C for 5 cycles (94°C for 30 sec, 48°C for 1 min, and 72°C for 1 min), followed by an increasing annealing temperature to 60°C in the second step for 35 cycles, and the cycle finished with a final extension of 72°C for 10 mins.

Amplicons were separated by electrophoresis on agarose gel and stained with RedSafe™ Nucleic Acid Staining Solution 20,000X (iNtRON Biotechnology, Korea). The gel was then visualised under ultraviolet light.

Table 1. Details of primers and their 5'-3' binding site sequences used in F1 hybridity assessment for the papaya ringspot virus-P resistance locus.

Primer	Primer sequences	Location	Reference
28.103-104	Forward: GTTGGACGCATCTACTCACG Reverse: CCTCTTCGACTACGCACACC	Supercontig 28 where *prsv-1* is located	Peace, personal communication
28.50	Forward: GCGTGCAACACTTTTCTCC Reverse: AAAAACTGCGTGAGATGTCG		
ctg-403	Forward: CCCCACCCATCATCCAACTA Reverse: CACTCATCAACCCACCCCAT	LG5	Chen et al., 2007
PBK41R	Forward: CCGCCCAAACTGCGGAACAC Reverse: CCGCCCAAACCCCCAACTAG	*prsv-1* locus	Dillon et al., 2006

2. *prsv-1* locus assessment

DNA of all the accessions previously mentioned (line 2.001, BC4#113, *V. pubescens* and *V. parviflora* and 18 lines of F1 PRSV-P) were assessed. The primer PBK41R (Table 1) was used to detect the *prsv-1* locus by generating the OPK4-1r SCAR marker (Dillon et al., 2006) using the PCR reaction conditions mentioned above. The OPK4-1r SCAR amplicons were subsequently digested by *Psi*I to generate *Psi*Ik4 CAPS markers (Dillon et al., 2006). Amplicons of 109 and 248 bp represented the dominant resistance allele and an amplicon of 372 bp represented the recessive allele at the *prsv-1* locus (Dillon et al., 2006).

3. Data analysis

The amplicons were subsequently scored manually for their presence (1) and absence (0) at each locus (Table 2). Comparison of the distinct banding patterns of the parents and their progenies determined the true hybrids and the non-hybrids.

Table 2. Scores for each of the alleles at four marker loci near the papaya ringspot virus resistance locus on linkage group 5 of four breeding accessions.

Marker	Location	Allele size	Allele score[1]			
			P1	P2	*pub*	*par*
28.103-104	SC28	320	0	1	0	0
		350	1	0	1	0
		400	1	0	0	1
28.50	SC28	380	1	0	1	0
		400	1	0	0	1
		450	0	1	0	0
		550	0	1	0	0
		800	1	0	0	1
ctg-403	LG5	180	1	0	1	1
		230	0	1	0	0
		400	0	1	1	0
OPK4-1r[2]	*prsv-1* locus	348	1	0	1	0
		372	1	0	0	1
*Psi*Ik4[3]	*prsv-1* locus	109	1	0	1	0
		248	1	0	1	0
		372	1	0	0	1

[1]1 = presence; 0 = absence of allele. P1 = BC4#113, P2 = *C. papaya* line 2.001, *pub* = *Vasconcellea pubescens*, *par* = *V. parviflora*.

[2]OPK4-1r is a SCAR marker resulting from amplification of PBK4rR primer (Dillon et al., 2006).

[3]*Psi*Ik4 is a CAPS marker resulting from the digestion of OPK4-1r amplicons with restriction enzyme *Psi*I (Dillon et al., 2006).

3. Morphological characterisation

Trees of the parental lines (BC4#113 and line 2.001) and the F1 family were grown to maturity and observed for their morphological characteristics at Lecker Farming, Mareeba, Australia. The observed characteristics included sex type, flower colour, fruit size and fruit shape.

RESULTS AND DISCUSSION

Intergeneric hybridisation

A total of 18 embryos were successfully established in vitro and developed into plantlets. The number of seeds yielded (2-20) was low compared with the general seed producing capacity of the female parents, which usually is in the range of 300-1000 seeds per fruit for line 2.001. Cross-incompatibility between *C. papaya* and *Vasconcellea* species was reported and this likely caused a reduction in seed production after intergeneric

hybridisation (Manshardt and Wenslaff, 1989). However, the influence from genetic distance may not have been the only factor limiting seed production. Environmental effects, including low rainfall, warm temperature and insufficient nitrogen, may also cause female sterility and have had an effect at this site previously (Garrett, 1995). Also, climate influences, such as low or excessive high temperature and low humidity adversely affect pollen quantity, viability and size (Allan, 1963). The number of viable seeds depended on the collection of viable pollen. Therefore, as the result was recorded in summer, this may also have contributed to the reduction of seed set in summer in north Queensland, where temperatures were usually above 35°C. Low fruit and seed set was also observed in some commercial cultivars during January and February 2017 (pers, obs.). Similar crosses in a backcrossing program were previously reported, with only 13 rescued embryos recovered (O'Brien and Drew, 2010).

F1 PRSV-P hybridity assessment

The amplification products from four primer pairs (Table 1) showed that all primers could amplify DNA of all parents and F1 progeny. All progeny possessed a DNA band pattern 100% homologous to those of both female and male parents (Figure 2), confirming true hybridity. The inheritance of the *prsv-1* locus from *V. pubescens* was confirmed using *Psi*Ik4 CAPS markers in all of the hybrid progenies (Figure 3). The expected bands at 108 and 240 bp were previously reported to be closely linked to the dominant resistance allele at *prsv-1* locus (Dillon et al., 2006) and these loci were detected in all the progeny. Interestingly, the presence of the OPK4_1r locus conflicted with the study of Dillon et al. (2006), who reported the absence of this locus in papaya, but agrees with that of O'Brien and Drew (2010), who reported the amplification of this locus in six papaya genotypes However, the amplification in those six genotypes only produced a 360 bp single band not able to be digested with *Psi*I (O'Brien and Drew, 2010), whereas multiple bands of 108, 240, 370, 1300, 1400 and 2,000 bp were detected in this study. This could be because different genotypes of *C. papaya* were used in this study than in the one by O'Brien and Drew (2010), which probably carried different alleles at this locus.

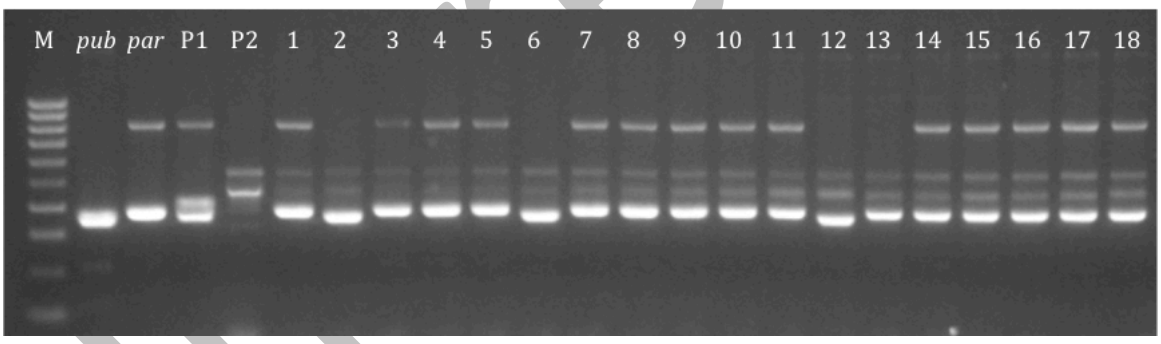

Figure 2. Simple sequence repeat marker amplified with the 28.50 primers. *pub* = *Vasconcellea pubescens*, *par* = *V. parviflora*, P1 = BC4#113, P2 = *Carica papaya* line 2.001 and 1-18 are F1 progeny.

DNA markers showed that all the progeny were true hybrids. Ten of the 18 F1s were successfully acclimatised and survived unde field conditions. The sex ratio was 1:1 (5 each of male and female), which agreed with the expected ratio when crossing between dioecious female and male papaya trees (Hofmeyr, 1938; Storey, 1938; Ming et al., 2007). Previous efforts to produce intergeneric hybrids from crosses between *C. papaya* and *V. pubescens* had resulted in female sex types only (Drew et al., 1998; O'Brien and Drew, 2009).

The F1 progeny had pink flowers both on female and male individuals (Figure 4). Fruit shape (blossom and tapered shape; IBPGR 1988) and size (8-12 g) in the F1 progeny

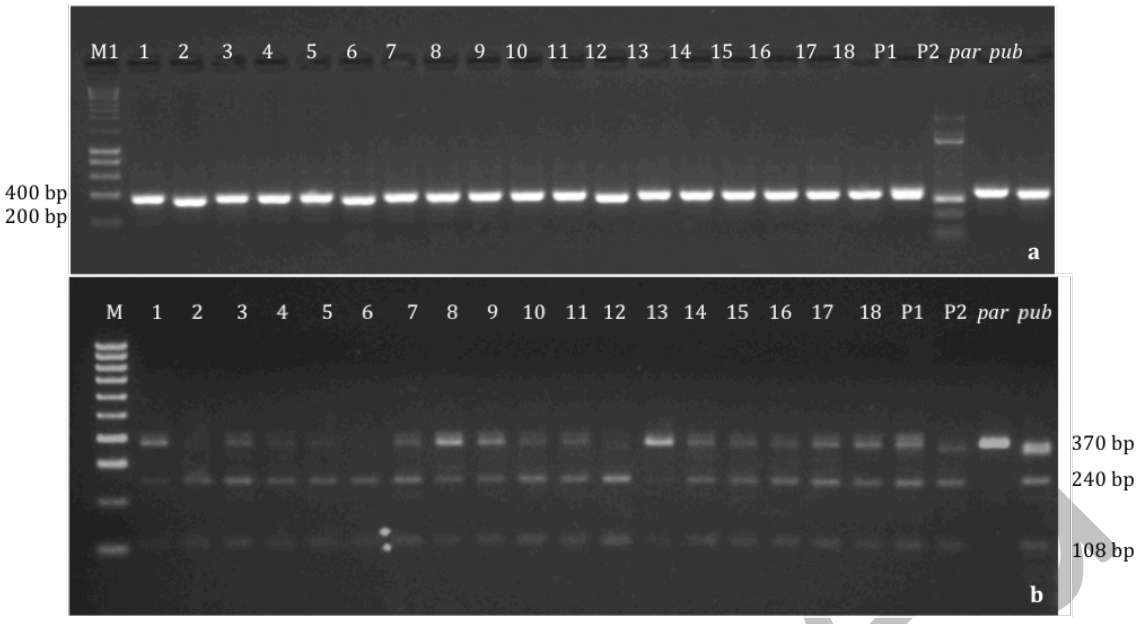

Figure 3.a) The OPK4-1r SCAR marker alleles amplified with the PBK41R primers; b) *Psi*I4k CAPS marker alleles generated from digestion reaction of OPK4-1r SCAR marker and restriction enzyme *Psi*I; lanes 1-18 are F1 progeny of *Carica papaya* line 2.001 x BC4#113; P1 = BC4#113; P2 = *C. papaya* line 2.001; *pub* = *Vasconcellea. pubescens*; *par* = *V. parviflora*.

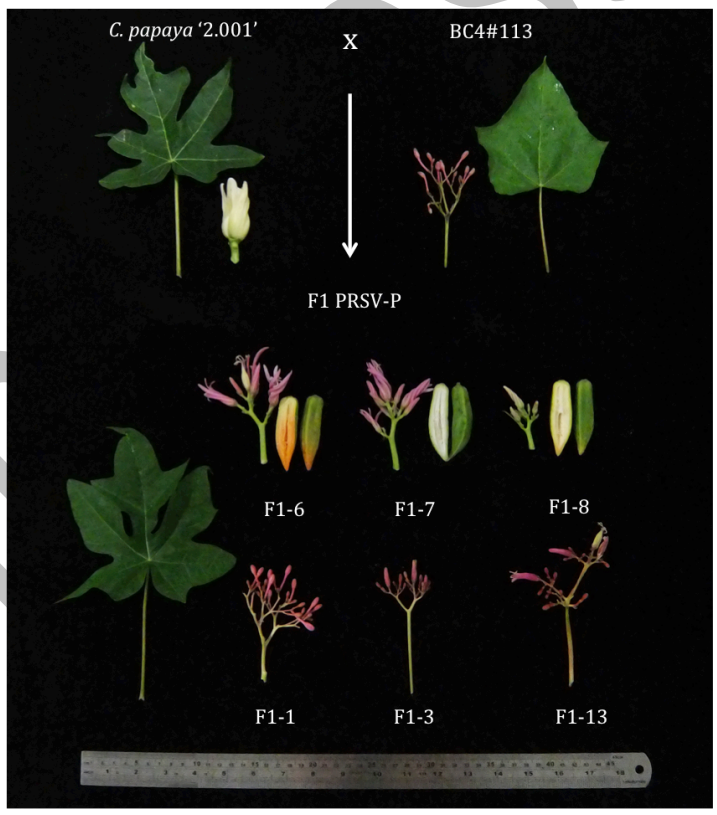

Figure 4. Leaves and flowers of parental lines (*Carica papaya* line 2.001 and BC4#113) and their F1 progenies. F1 lines 6, 7 and 8 show female flowers and fruits, and lines 1, 3 and 13 show male flowers.

were also inherited from *V. parviflora*. This result re-confirmed the hybridity test by DNA markers. The hypothesis of using *V. parviflora* as a bridging species for intergeneric hybridisation between *C. papaya* and *Vasconcellea* spp. was first proposed by O'Brien and Drew (2009) and this is the first report to confirm the success of this approach.

CONCLUSION

The intergeneric hybridisation between *C. papaya* and BC4#113 of *V. pubescens* x *V. parviflora* was established and embryo rescue was performed to produce 18 F1 lines, which were successfully micropropagated and acclimatised. This is the first report to present successful intergeneric hybridisation between *C. papaya* and *Vasconcellea* spp. using *V. parviflora* as bridging species. Based on both DNA markers and morphological traits, these progenies were true hybrids and contained the *prsv-1* locus. These F1 will be tested for resistance to PRSV-P and if resistant, will be used as a donor to transfer the resistance allele to cultivated papaya.

ACKNOWLEDGEMENTS

This project is funded by Horticulture Innovation Australia Limited using the papaya industry levy and funds from the Australian Government (PP15000: New genetic targets to improve quality in papaya).

Literature cited

Allan, P. (1963). Pollen studies in *Carica papaya*. II: Germination and storage of pollen. SA J Agric. Sci. *6*, 613-624.

Bau, H.J., Cheng, Y.H., Yu, T.A., Yang, J.S., and Yeh, S.D. (2003). Broad-spectrum resistance to different geographic strains of papaya ringspot virus in coat protein gene transgenic papaya. Phytopathology *93*, 112-120.

Bau, H.J., Cheng, Y.H., Yu, T.A., Yang, J.S., Liou, P.C., Hsiao, C.H., Lin, C.Y., and Yeh, S.D. (2004). Field evaluation of transgenic papaya lines carrying the coat protein gene of papaya ringspot virus in Taiwan. Plant Dis. *88*, 594-599.

Chen, C., Yu Q., Hou, S., Li, Y., Eustice, M., Skelton, R.L., Veatch, O., Herdes, R.E., Diebold, L., Saw, J., et al. (2007). Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. Genetics *177* (*4*), 2481-2491.

d'Eeckenbrugge, G.C., Drew, R.A., Kyndt, T., and Scheldeman, X. (2014). *Vasconcellea* for papaya improvement. In Genetics and Genomics of Papaya, R. Ming and P.H. Moore, eds. (New York, NY, USA: Springer), 10: 47-79

Dellaporta, S.L., Wood, J., and Hicks, J.B. (1983). A plant DNA minipreparation: Version II. Plant Molec. Biol. Rep. *1*, 19-21.

Dillon, S., Ramage, C., Ashmore, S., and Drew, R.A. (2006). Development of a codominant CAPS marker linked to PRSV-P resistance in highland papaya. Theor. Appl. Genet. *113*, 1159-1169.

Drew, R.A. (1988). Rapid clonal propagation of papaya *in vitro* from mature field-grown trees. HortScience, *23*, 609-611.

Drew, R.A. (1992). Improved techniques for *in vitro* propagation and germplasm storage of papaya. HortScience, *27*, 1122-1124.

Drew, R.A., O'Brien, C.M., and Magdalita, P.M. (1998). Development of *Carica* interspecific hybrids. Acta Hortic. *461*, 285-292.

Drew, R.A., Persley, D.M., O'Brien, C., and Bateson, M.F. (2005). Papaya ringspot virus in Australia and the development of virus resistant plant. Acta Hortic. *692*, 101-106.

Drew RA, Smith NG (1986) Growth of apical and lateral buds of papaw (*Carica papaya* L.) as affected by nutritional and hormonal factors. J. Hortic. Sci. *61*, 535–543.

Fitch, M.M., Manshardt, R.M., Gonsalves, D., and Slightom, J.L. (1993). Transgenic papaya plants from agrobacterium-mediated transformation of somatic embryos. Plant Cell Rep. *12*, 245-249.

Garrett, A. (1995). The pollination biology of papaw (*Carica papaya* L.) in central Queensland, PhD Thesis, (Rockhampton, Australia: Central Queensland University), pp.170.

Gonsalves, D. (1998). Control of papaya ringspot virus in papaya: A case study. Ann. Rev. Phytopath. *36,* 415-437.

Hofmeyr, J.D.J. (1938). Genetical studies of *Carica papaya* L. I. The inheritance and relation of sex and certain plant characteristics. II. Sex reversal and sex forms. SA Dept. Agri. Sci. Bull., *187*, 64.

International Board for Plant Genetic Resources (IBPGR) (1988). Descriptors for Papaya, IBPGR, Rome.

Kanchana-udomkan, C., Razali, R.M., Drew, R., and Peace, C. (2015). Developing DNA markers for a wild source of resistance to Papaya Ringspot Virus. Acta Hortic. *1100*, 47–52.

Kung, Y.J., Bau, H.J., Wu, Y.L., Huang, C.H., and Chen, T.M., and Yeh, S.D. (2009). Generation of transgenic papaya with double resistance to papaya ringspot virus and papaya leaf-distortion mosaic virus. Phytopathology *99*, 1312-1320.

Magdalita, P.M., Drew, R.A., Adkins, S.W., and Godwin, I.D. (1997). Morphological, molecular and cytological analyses of *Carica papaya* × *C. cauliflora* interspecific hybrids. Theor. Appl. Genet. *95*, 224-229.

Manshardt, R.M., and Wenslaff, T.F. (1989). Interspecific hybridization of papaya with other *Carica* species. J. Amer. Soc. Hort. Sci. *114*, 689-694.

Ming, R., Yu, Q., and Moore, P.H. (2007). Sex determination in papaya. Semin. Cell Developm. Biol. *18*, 401-408.

O'Brien, C.M., and Drew, R.A. (2009). Potential for using *Vasconcellea parviflora* as a bridging species in intergeneric hybridisation between *V. pubescens* and *Carica papaya*. Austr. J. Bot. *57*, 592-601.

O'Brien, C.M., and Drew, R.A. (2010). Marker-assisted hybridisation and backcrossing between *Vasconcellea* species and *Carica papaya* for PRSV-P resistance. Acta Hortic. *859*, 361-368.

Purcifull, D.E., Edwardson J.R., Hiebert E., Gonsalves D. (1984) Papaya ringspot virus. CM1/AAB descriptions of plant viruses NO. 292 (Wallingford, UK: CAB International), pp. 8.

Razean Haireen, M.R. (2013). Identification, characterisation and expression of PRSV-P resistance genes in *Carica* and *Vasconcellea*. PhD Thesis (Brisbane, Australia: Griffith University), pp.224.

Sharma, S.K., and Tripathi, S. (2013). Overcoming limitations of resistance breeding in *Carica papaya* L. In Plant Virus-Host Interaction: Molecular Approaches and Viral Evolution, R.K. Gaur, H. Thomas, and S. Pradeep, eds. (Waltham, USA: Academic Press), p.430.

Storey, W.B. (1938). Segregation of sex types in solo papaya and their application to the selection of seed. Proc. Amer. Soc. Hortic. Sci. *35*, 83-85.

Thomas, J.E., and Dodman, R.L. (1993). The first record of papaya ringspot virus-type P from Australia. Austral. Plant Pathol. *22*, 2-7.

Tripathi, S., Ferreira, S.A., and Gonsalves, D. (2008). Papaya ringspot virus-P: characteristics, pathogenicity, sequence variability and control. Molec. Plant Pathol. *9*, 269-280.

Appendix 2: List of SSR primers used for genotyping the diverse germplasm collection

Oligo Name	Sequence 5' to 3'		Tºm	Motiff	Number of allele	Size range (bp)
CT9395	F:	CGTGACGGTGAGATATCGTGA	66.4	(TA)11	3	948-992
	R	GCGACATGTGGAGTGGAGAT	66.2			
CT9835	F:	GGACGTGTCTGGAGCATACT	66.2	(AT)13	2	251-529
	R	TATCCAATGTGGCGCACTGT	67.3			
CT80-2	F:	GCTTTGCCACTTGACACCAG	64	(TA)14(AT)13	2	622-800
	R	TCTCAAGCTTTGCACACCCA	63.3			
CT241	F:	TACCCACCTTTACCTCCGTG	62.5	(TA)17	4	560-603
	R	TGTTGGTAAGATTCAGATTGCAG	61.9			
CT204	F:	ACGTATCCAGACAGTCCCAA	64	(A)10(AAC)5(AT)17(A)16	2	170-203
	R	ATGGTCCCTGTACTCCTCCT	65.2			
CT0	F:	TCACTGATTCATTTTCCGGCA	67.8	(AT)9(TGTA)5	2	949-1034
	R	ATATGTGGTGTTTGGGGGCA	67.4			
CT59	F:	ATGTGAAAATTGTTGGAGACCC	57.13	(TA)19	2	997-1035
	R	TTTGAAGGGTAAACAGCGGC	59.04			
CT173-1	F:	ACCGTAAATGAGAGGTTGTGAT	57.45	(CT)25(AT)8	2	201-232
	R	GTGGAGTAGGTTTGTTCCCC	57.8			
CT28	F:	TCCAGGAATCAACCCTTGCT	58.92	(T)17(CA)7g(AT)17g(TA)6	2	233-243
	R	ACAACTTAACATGGCAGAGCA	58.14			
CT184	F:	GCTTGCCCCTTTATACACACA	58.56	(AT)8(TATG)5	2	844-1000
	R	TCTGTTCGTATGCCACCTTTG	58.58			
CT149	F:	ATGCTTCAACGGCCATAGTT	63.7	(AG)26	2	480-512
	R:	AGGTAGCAACAAGGTTTCATCAG	63.6			
CT209-2	F:	TTGTGCTTGGAAAATGCCCA	68.8	(CAT)5	2	435-477
	R:	AGTTCTTCCGTTGGTGGGTA	63.7			
CT23-2	F:	GACATCTCCGAATACAGCGC	65.1	(TG)6(AC)6	2	466-541
	R:	CTTCAACGCCAAACCAGGAA	67.5			
CT3-5	F	AGGTCAGTGTCGCCAAGTTA	62.8	(AAT)14(T)10	2	282-372
	R	GTTCTCGTGCAACCGATACC	64.8			
CT32	F:	GCATTCCCAAAATCTGCTGG	66.9	(AT)16	2	369-409
	R:	GGAGCTACAACAAAAGGGCA	64.4			
CT36	F:	GCTTCCATTATTGCACTTGTTGG	66.6	(CT)33	2	393-862
	R:	ATCCCACCACAAGCCTCATC	66.6			
CT46-2	F:	GGCAGGCATTGTAGTAACCG	64.6	(TA)15	2	551-610
	R:	GGAACTGGGTGTAAAGGAAAACT	63.5			
CT50	R:	TCAAGTTCGATGTGCTCAGC	64.7	(A)11(AT)10	2	370-454
	F:	AATGGCAGTCTCAGTTACCG	62.2			
CT64	F	ACTTGCTTATCAGTGCCGCT	64.2	(AG)25	2	596-679
	R	GAACTTTTCTTCTTCCGCATGT	63.7			
CT80	F:	CAATTCGCTGATGGTGTCTGA	66.6	(TA)12(TA)17	2	223-281
	R:	CACCCATGTACTGCTGCAAA	65.2			
CT882	F	AGCTAACCAGTTTAATACCGCT	60.2	(TA)13	2	348-403
	R	TTACTTGGTTGTGGGGTCGG	67.2			
CT104-F	F	AGGCAGAAATGGGAGACACG	66.8	(A)15(T)12g(A)12	2	318-361
	R	GACCCCAATCACGAAACGATA	66			
CT110-F	F	AGTTTCGAGTCAAATGGAGGG	64.5	(C)15	4	387-463
	R	TGGGCTTCATCTTGGTCCAT	67			
CT126-F	F	AGCTTCTCATAACCTTAGCCAA	61.2	(CAG)11	2	399-616
	R	CCGATTGCAGACTTTGGACC	67.1			

Oligo Name	Sequence 5' to 3'		Tºm	Motiff	Number of allele	Size range (bp)
CT149-2F	F	ATGCTTCAACGGCCATAGTT	63.7	(AG)12	2	480-518
	R	AGGTAGCAACAAGGTTTCATCAG	63.6			
CT150-F	F	GGATTTTGATACCGTTCATACCC	64.3	(TA)8(AAT)6	2	431-525
	R	CACTTTCCCAAACGAAGCCTAA	66.3			
CT151-F	F	ATGTGTCTACTTCGGGCTCC	63.5	(TTAT)5(A)10	2	437-472
	R	AGAACAAAACACCCAACCGT	63.6			
CT152-F	F	TGGCGCTTTGTTAAATTGGC	67.1	(T)17(A)11	2	276-345
	R	ACATGGTGTAAGAAGCGAATAAC	61.3			
CT161-F	F	CCCTTTTCATTTTCCGTTTCGTA	66.8	(AC)7(AT)7	2	359-406
	R	TTTTACTGCGCCTACTAACGA	61.8			
CT2-2F	F	ACTCTTGGAAGTCATATAGGCAA	61	(T)10(AT)10	2	500-718
	R	CGTTTCTAGTTCCTGCATGTG	62.7			
CT215-F	F	TCACTTAATCCCTTGGCAAAGT	63.9	(TAT)6	2	625-709
	R	AAAATGCATGACACCTGTACAAA	63.4			
CT22-F	F	GGCTTCCATAGAGATCAAAGTGG	65.1	(TA)6(AG)7	3	287-358
	R	ACCCTGTTTAGATTCCTTTCGG	64.4			
CT223-F	F	TATCATCTAGCACCGCCACA	64.6	(A)13(AAC)5	3	600-684
	R	GTGGAACGGGGCTTGAAATT	67.3			
CT242-F	F	GGAATTTATGCACAATCTGGACA	65	(TTA)5(TAA)6	2	282-377
	R	TGTTGCTAGCTGGTATACTTGAG	60.5			
CT270-F	F	GACATCCATTCAAGACTTCCCC	66	(AC)9(AT)6	2	284-316
	R	TCATTCTCGCTCATCTTAAGGAA	64.2			
CT28-2F	F	TGGGGTCCAGGCTTATTTGA	66.7	(T)17(CA)7g(AT)17g(TA)6	2	391-453
	R	ACAGCACCGACCTTTTCTTC	63.6			
CT286-F	F	TTGAAACCATGTGCTCCGTC	66.4	(TTAA)6	3	399-479
	R	ATGAATAGCCGCTGGAGGG	66.6			
CT32-2F	F	GGTTTGTAGAGAGGAGGGTGA	62.5	(AT)16	3	761-877
	R	TGGAAGGGTGGTATTTTGATGC	67			
CT33-F	F	TCTTGTGGGCTCGCTCTTAT	64.2	(T)16(T)10(A)11	2	498-520
	R	ACAAAGACTCAAACGATGCTGT	63			
CT335-F	F	ATGCCTCCAATTTGCTTTGC	66.4	(TG)5	2	547-587
	R	GCACACATAGGTCCAGATTTGA	64.2			
CT361-F	F	ACAAGCCAAGAACTCAAACAAGT	63.2	(A)16(TTA)5(A)10	2	525-588
	R	ATTGCCCTCCCTCCCAAAAT	67.8			
CT37-F	F	TGTTTGGGTGAGTAAAGAAACGT	63.8	(TAT)8(C)20	2	190-233
	R	TGGAGTGTATGACCTTTCCATG	64			
CT40-F	F	GGACTGACACAAACAATGACCA	65.2	(TC)13(CTT)5	2	588-672
	R	TGGGAGTTTTGTTCATGGGG	67.2			
CT449-F	F	GAGGCATTTACCACATGCTTTT	64.2	(T)14(CT)11	2	481-510
	R	CGTCCAAGGAGTGAGGTTCT	63.7			
CT47-F	F	ACTGCAATGAGAATCAATCGC	64.4	(AAG)5(A)15	2	162-195
	R	GTGACCAACGCTCAATTCCT	64.5			
CT524-F	F:	TAGTTGGGATAGGTGGCACA	63.4	(AG)12	2	246-276
	R:	TGAGTACAGCGGTGAATTACTTG	63.3			
CT532-F	F	GTGGAGGTTCTTGCAGACG	64	(AT)25	2	439-486
	R	TGAACACGTTTTGGTATTGCG	65.9			
CT74-F	F	TGGGTGGGATGATTTTGGGA	69.7	(T)20	2	385-411
	R	TCTGCTTACTACCCATGCCC	64.3			

Appendix 4: Database of papaya germplasm

Traits	GR1-C	GR1-T	GR2-C	GR2-T	GR3-C	GR3-T	GR4-C	GR4-T	GR5-T	GR7_C	GR8-T	GR9-C	GR9-T	GR10-C	GR10-T	GR21-T	GR23-C	GR23-T	GR24-T	GR25-C	GY26-C	GY26-T	GY27-T	GY29-C	GY29-T	HD1-T	HD2-T	R2-T	R3-T	R4-T
	RB1		RB2		RB4		Solo		inda Sol	Holland	aeg Nu	First Lady		TN#2		-Dimbul	TS2		P-Babind	Paris	H1B		H11B	H29		ey drew	ey Drew	olo 30	Red #3	Solo 8.2
Trunk circumference	48.00	49.20	45.67	46.40	50.00	42.80	46.67	42.00	34.20	33.50	36.80	40.00	40.00	44.00	40.00	42.00	39.75	37.20	40.50	36.00	38.00	55.40	45.40	37.00	37.40	31.33	43.67	57.00	41.67	41.40
Peduncle length	2.40	1.20	3.00	2.00	2.00	1.80	3.00	1.00	1.00	3.00	3.00	1.50	1.00	3.00	2.80	1.50	2.25	2.60	1.50	3.00	2.67	2.80	3.00	3.00	2.00	1.00	1.00	1.00	1.67	1.00
Fruit length	17.80	19.00	19.00	20.30	14.68	14.94	12.47	13.00	14.35	27.50	37.60	11.50	16.98	28.53	22.50	20.00	14.05	12.90	17.70	19.80	18.60	17.10	16.13	15.30	14.55	13.23	13.67	14.57	19.07	13.26
Fruit width	10.26	9.78	12.00	11.84	8.30	9.04	8.27	8.00	7.68	9.50	12.85	6.80	9.80	11.10	9.23	9.00	7.90	10.33	13.60	15.50	11.47	12.00	10.45	12.83	12.75	11.07	12.10	7.33	9.80	6.76
Fruit weight	909.40	868.60	####	####	555.00	655.40	437.67	416.67	367.00	####	####	451.00	###	1549.00	888.75	849.00	423.50	533.33	1518.00	1195.00	1562.67	###	899.00	1407.00	1300.25	803.33	934.67	###	925.00	####
Number of fruit	13.20	23.40	11.00	16.80	19.00	28.00	23.67	29.20	24.20	6.00	24.80	3.00	25.00	11.00	23.00	26.75	16.75	28.20	18.75	11.35	12.67	19.40	13.60	15.33	14.80	16.33	10.00	46.00	22.67	36.20
Number of carpelloid fruit	0.40	2.20	2.33	1.60	2.50	2.40	4.67	0.60	2.60	0.75	0.20	0.00	2.00	2.67	1.40	2.75	8.75	10.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.33	0.33	4.00
Yield gap	0.60	22.00	0.67	22.00	0.75	4.00	0.00	2.00	2.00	2.50	10.00	3.00	6.00	1.00	12.00	15.00	1.25	4.00	27.50	2.25	1.67	24.00	32.00	0.67	20.00	33.33	30.00	0.00	3.33	0.00
Skin gloss	2.20	3.00	4.00	4.00	3.00	3.40	1.00	2.00	1.25	1.00	3.00	2.00	1.75	2.33	2.25	1.00	1.00	2.00	2.00	4.00	2.00	2.60	1.00	2.00	1.25	1.33	1.33	1.00	2.33	1.00
Skin freckle	2.40	4.00	1.33	1.00	1.00	1.40	11.00	4.00	4.75	1.00	3.50	3.00	4.50	3.00	2.25	3.00	4.00	1.00	4.00	2.00	1.33	1.20	2.25	1.33	2.00	5.00	5.00	3.00	4.00	1.00
Skin colour	6.60	6.20	7.00	7.80	7.00	7.00	6.33	7.00	6.50	8.00	7.00	7.00	7.00	7.00	7.00	6.67	6.75	8.00	7.00	7.00	7.00	6.60	6.25	6.33	6.00	6.33	7.00	6.00	6.33	6.60
Fruit firmness	2.20	2.00	3.00	3.00	2.00	3.00	1.00	1.67	2.25	3.00	3.00	1.00	1.75	1.67	1.75	2.33	2.25	2.00	3.00	3.00	2.67	1.80	2.75	2.00	2.75	1.00	1.00	1.67	2.33	1.60
Fruit shape	6.80	7.00	7.00	7.00	6.00	7.00	5.00	6.00	6.00	9.00	9.00	8.00	6.75	9.00	7.00	6.67	6.75	3.33	5.00	5.00	4.00	4.00	4.00	2.67	2.00	5.00	5.00	7.00	7.00	6.00
Teat shape	2.80	3.00	1.33	3.00	2.00	2.80	2.67	2.00	2.50	4.00	4.00	4.00	2.25	4.00	4.00	2.33	3.25	3.00	3.00	4.00	3.00	3.00	2.75	3.00	3.00	4.00	4.00	3.00	3.00	3.00
Stalk insertion	2.00	2.00	2.00	2.00	3.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.67	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.25	1.00	1.00	2.00	2.00	2.00
Cavity length	12.66	13.16	12.87	14.90	9.83	10.16	8.77	8.00	9.45	21.78	33.00	13.00	12.33	20.37	15.75	13.33	10.40	9.17	13.00	13.60	13.70	11.80	11.63	10.70	9.63	8.67	9.17	9.80	13.43	8.30
Cavity width	5.56	4.76	5.07	5.76	3.78	4.60	4.23	4.53	4.38	4.18	7.10	2.50	5.35	5.87	4.55	4.83	3.38	6.93	9.00	9.20	5.57	20.20	5.45	6.93	7.50	7.13	7.67	3.67	4.77	3.00
Cavity shape	2.40	2.40	3.33	2.20	3.25	3.25	2.67	3.33	4.00	1.00	1.00	1.00	2.75	1.00	1.00	1.67	2.00	1.67	1.00	2.00	2.33	2.00	2.25	2.00	2.00	2.00	2.00	4.00	1.00	4.00
Flesh colour	7.80	7.00	8.00	8.00	7.00	7.00	7.67	6.00	6.00	7.25	7.00	6.00	7.25	7.33	7.75	6.67	7.00	8.00	5.00	5.00	4.67	5.00	5.00	5.33	5.00	4.67	5.00	6.00	6.00	6.00
Consistency in flesh colour	3.00	3.00	2.67	3.00	3.00	3.00	3.00	3.00	2.75	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
Flesh thickness	2.34	2.22	3.37	2.80	2.03	1.98	1.67	1.63	1.50	2.03	2.25	22.00	2.00	2.60	2.00	1.93	1.93	1.43	2.00	3.20	3.37	2.70	2.58	3.13	2.30	1.80	2.00	1.63	2.13	1.60
Flesh texture	1.00	1.20	3.00	3.00	1.25	2.60	1.00	1.00	1.50	3.00	3.00	1.00	1.00	1.00	1.50	2.67	2.75	2.33	2.00	1.00	2.33	2.60	2.50	1.33	2.20	1.00	1.00	1.00	1.00	1.00
Flesh sweetness (°Brix)	10.10	11.88	10.07	9.82	11.08	11.88	12.70	13.00	12.40	9.63	10.20	11.40	12.73	11.63	12.13	11.20	10.05	15.27	11.00	10.70	9.33	9.24	10.30	8.30	11.78	11.57	12.37	12.93	12.40	12.78

Appendix 5: DNA profile of papaya germplasm

Oligo Name	Allelle	DNA sample																	
		1	2		3	4	5	6	8	9	10	21	22	23	24	25	26	28	29
		RB1	RB2	SKB	RB4	Solo	Linda Solo	Holland	Kheg Nuan	First Lady	Tainung #2	OP-Dimblah	JC2	TS2	OP-Babinda	OP-Paris	H1B	H13	H29
CT9395	1	0	0	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	1
	2	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0
CT9835	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1
	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
CT80-2	1	1	1	0	0	0	0	1	1	1	1	1	0	1	1	1	1	0	1
	2	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	0
CT241	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1
CT204	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
CT0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1
	2	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1
	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1
	4	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
CT59	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1
	2	0	0	1	1	1	0	0	0	1	0	0	0	1	0	1	1	0	0
CT173-1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1
CT28	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CT184	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	0
CT149	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1
	2	0	0	0	1	1	1	0	0	1	0	0	1	1	1	0	1	1	1
CT209-2	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1
	1	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1
CT23-2	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	0	0	0
	2	0	0	0	1	1	1	0	0	1	1	0	1	1	0	0	1	1	1
CT3-5	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1
	2	0	0	1	1	1	1	0	0	1	0	0	1	0	0	0	0	1	0
CT32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
	2	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1
CT36	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0
CT46-2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CT50	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CT64	1	0	0	1	1	0	1	1	1	0	0	1	1	1	0	1	0	0	0
CT80	1	0	0	1	1	0	1	1	1	0	1	1	1	1	1	1	0	0	0
	2	0	0	1	1	1	1	0	0	1	0	0	0	1	0	1	1	1	1
CT882	1	0	0	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	0
	2	0	0	1	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1
CT104	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	2	0	0	1	0	1	1	0	0	1	1	1	1	1	1	1	0	0	0
CT110	1	0	0	1	1	0	1	1	1	0	0	1	1	1	0	1	0	0	0
	2	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
CT126	1	0	0	1	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1
	2	0	0	0	1	1	1	0	0	1	0	0	1	1	1	0	1	1	1
CT149-2	1	0	0	0	1	0	1	1	1	0	0	1	1	1	0	0	0	0	1
	2	0	0	1	1	1	0	0	0	0	0	1	1	0	0	1	0	0	0
	3	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
CT150	1	0	0	1	1	0	1	1	1	0	0	1	0	0	1	0	0	0	0
	2	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
CT151	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1
	2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1
CT152	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0
CT161	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1
	2	0	0	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1
CT2-2	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	0
	2	0	0	1	1	1	1	0	0	0	1	0	0	0	0	1	1	1	1
CT215	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	2	0	0	1	1	1	0	0	0	0	0	0	0	1	1	1	1	0	1
CT22	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1
	2	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	1	1
CT223	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	0	0	1
	2	0	0	1	1	1	1	0	0	1	1	0	1	1	0	1	1	1	1
CT242	1	1	1	0	0	0	0	1	1	0	0	1	0	0	1	1	0	0	0
	2	0	0	1	1	1	1	0	0	1	1	0	1	1	0	0	1	1	1
CT270	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	0
	2	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0	1	1	1
CT28-2	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	1	0	0	0
	2	0	0	1	1	1	1	0	0	1	0	0	0	1	1	1	1	1	1
CT286	1	1	1	0	1	0	1	1	1	0	0	1	0	0	0	0	0	0	0
CT32-2	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	0
	2	0	0	1	1	1	0	0	0	1	1	0	1	1	1	1	1	1	1
CT33	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	0
	2	0	0	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1
CT335	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	2	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0
CT361	1	0	1	1	1	0	1	1	1	0	0	1	1	1	1	1	0	0	1
	2	0	0	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1
CT37	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1
	2	0	0	0	1	1	1	0	0	1	1	0	1	1	0	1	0	1	0
CT40	1	1	1	1	1	0	1	1	1	0	0	1	1	1	0	1	0	0	0
	2	0	0	1	1	1	0	0	0	1	0	0	0	1	1	1	0	0	0
CT449	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1
	2	0	0	1	1	1	1	0	0	0	0	0	1	0	1	0	0	1	0
CT47	1	0	0	1	0	1	1	0	0	0	0	0	1	1	0	0	1	0	0
	2	1	1	1	1	0	1	1	1	0	0	1	1	1	1	0	1	0	0
	3	0	0	1	0	1	1	0	0	1	1	0	1	1	0	1	1	1	1
CT524	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	0	0	1
CT532	1	1	1	1	1	0	1	1	1	0	0	1	1	1	0	1	0	1	0
	2	0	0	0	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1
CT74	1	1	1	1	1	0	1	1	1	0	0	1	1	0	0	1	1	1	0
	2	0	0	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1

DNA Finger-printing of Commercial Australian Papaya Cultivars by Genome-by-Sequencing Analysis

Mai Nantawan

6 July 2018

Study Description: DNA profile documentation for Australian papaya cultivars

Experimental Design

Seven papaya cultivars derived from Australian germplasm, *Carica papaya*, (1B, RB1, RB2, RB4, Sunrise Solo, Sekaki, ER62) were included in this study. Total of 45 individuals were genotyped to assess genetic diversity among and within cultivars.

Genotyping

The 7 papaya cultivars were subject to genotyping by sequencing (GBS) analysis (3-4 replicates for each cultivar (n=45)). The DNA were sent for genotyping-by-sequencing (GBS; AGRF, Adelaide). The AGRF uses ddRAD based library preparation protocol, as described in Peterson *et al.* (2012). The DNA was digested using a combination of restriction enzymes (*PstI* and *MseI*) and only tags with both RE sites (one in each end) were selected for library preparation and sequenced on 2 lanes of an Illumina HiSeq2500 sequencing platform, producing 100 bp single-end reads. Basic bioinformatics analysis was performed by AGRF, mainly running the Stacks v1.46 pipeline (Catchen *et al.* 2011; Catchen *et al.* 2013). In brief, the raw reads were demultiplexed, and barcodes and adaptors were removed by Stacks process_radtags.pl. The software takes sequences in FASTQ.GZ format as an input and deconvolutes each read according to the inline barcodes. The pipeline also checks for read quality and restriction site presence. It creates a separate FASTQ file for each sample. It then automatically trims FASTQ files to the size of the shortest read minus 2 bases to compensate for differences in read length due to any variation in barcode sequences. A catalogue of consensus tag sequences was built for each sample (. fq.gz file).

Analysis Steps

1. Assess quality and clean reads from GBS: FastQC and Trim Galore
2. Map reads to the draft papaya genome (Ming et al. 2008): using Bowtie2
3. Perform Quality control BAM files using Picard.
4. Call variants from the mapped reads with Mpileup and BCFtools (Li et al. 2009, 2011)
5. Filter resulting .vcf file to remove missing ALT loci (using simple expression c5=="." and SnpSift Filter) and only contain high quality SNPs with DP>5 and QUAL >20)
6. Filter for common variant genotype (position) among cultivar population: using SnpSift (Cingolani et al. 2012) (countHom() >=) | (countHet() >=)
7. Filter out for specific variant genotype (position) for each cultivar using VCFtool (VCFtool intersect multiple VCF dataset).
8. Filter out resulting specific variant genotype (position) for each cultivar that located in gene interval (not on gene regions: BED, gff file from Plaza: annotation.cpa.gff) using SnpSift (SnpSift interval Filter variants)
9. Merge filtered files from each cultivar for further analysis and data visualisation using VCFtools (Merge multiple datasets)
10. Generate heatMap and graphic using R script (vcfR package)

Methodologies

- **Quality control: Assess read quality and clean reads**

Data (raw reads (.fq.gz file) and reference genome (.fa file)) was imported to Galaxy platform.

FastQC: Run **FastQC Read Quality reports.**

Trim Galore: Run **Trim Galore! Quality and adapter trimmer of reads** on the imported FastQ file.

FastQC: Re-run **FastQC Read Quality reports** to check the impact of Trim Galore

- **Mapping to a reference genome**

Trimmed reads were mapped to the *C. papaya* reference genome, 'SunUp' variant (Ming *et al.* 2008) using Bowtie2 (version 2.3.2) with the --very-sensitive option in --end-to-end (global) mode.

Mapping: Bowtie2: Search in the tool bar on the left the mapper 'bowtie2' and run the mapper with the dataset.

Select "use a genome from history and build index"

Set read group to "Picard style" with –very sensitive end to end mode.

- **Quality control Mapped reads**

Picard: MarkDuplicatesWithMateCigar: examine aligned records in resulting BAM files to locate duplicate molecules (download BAM file to hard drive)

Picard: CleanSam: perform Sam/Bam grooming

- **Variant calling *on Metadata (all cultivars)***

Mpileup: generate variant from BAM file output format in vcf file (document ID# 1224) **BCFtools:** to call variant from vcf file Pval threshold 0.5

Filter data on any column using simple expressions (Galaxy Version 1.1.0): to Filter missing ALT loci (".") out

BCFtools: to filter high quality SNP with DP>5, and GQ>20 result high quality SNP/INDEL variants

SnpSift interval Filter variants **(Snpeff tool in Galaxy):** to select variants in VCF file that locate in gene region (from annotation.cpa.gff Plaza website database) result significant SNP/INDEL variant close to genes.

- **Genetic diversity analysis**

Use R script for inspection remaining SNP (VCF file) convert to genlight object and draw dendrogram based on the UPGMA algorithm, with 100 bootstrap replicates

```
# Set working directory and load R packages
setwd("~/Desktop/R-session/GBS_germplasm_data_analysis")
# download packages for VCF inspection
install.packages("vcfR")
library(vcfR)
install.packages("adegenet")
library(adegenet)
```

```
install.packages("adegraphics")
library(adegraphics)
install.packages("pegas")
library(pegas)
install.packages("StAMPP")
library(StAMPP)
install.packages("lattice")
library(lattice)
install.packages("gplots")
library(gplots)
install.packages("gape")
library(ape)
install.packages("ggmap")
library(ggmap)
install.packages("poppr")
library(poppr)
install.packages("ape")
library(ape)
install.packages("RColorBrewer")
library(RColorBrewer)
install.packages("vcftools")
library(vcftools)
install.packages("ggplot2")
library(ggplot2)
install.packages("dartR")
library(dartR)
##########################################################################
#
# Download Vcf file from Galaxy platform (document# 1225) and check data quality (VCF object)
vcf <- read.vcfR ("NEWGalaxy1251-[bcftools_annotate_on_data_1248].vcf")
head(vcf)  #check VCF
vcf@fix[1:10,1:5]

# Plot statistic summary
chrom <- create.chromR(name='RAD_data', vcf=vcf)
```

```
plot(chrom)


# Check read depth distribution per individual and plot

dp <- extract.gt(vcf, element='DP', as.numeric=TRUE) pdf("DP_RAD_data_zoom.pdf", width = 10, height=3)

par(mar=c(8,4,1,1))

boxplot(dp, las=3, col=c("#C0C0C0", "#808080"), ylab="Read Depth (DP)",las=2, cex=0.4, cex.axis=0.5, ylim=c(0,50))

abline(h=8, col="red")

dev.off()


# Load population data (.txt file: show detail of population/groups)

pop.data <- read.table("NEWpopulation_data_gbs.txt", sep = "\t", header = TRUE)

all(colnames(vcf@gt)[-1] == pop.data$AccessID)


# Convert to genlight object for further analysis and specify polypoid as 2 (diploid)

aa.genlight <- vcfR2genlight(vcf)

ploidy(aa.genlight) <- 2


# Add population data to genlight object (etc. cultivar) for further analysis

pop(aa.genlight) <- pop.data$Cultivar

aa.genlight


#############################################################
#

#Population genetic analysis (metadata)

#1.Distance metrix can use function dist( ) or bitwise.dist() or as.matrix()

aa.genlight.dist <- poppr::bitwise.dist(aa.genlight)

#******************


#2.distance tree reconstruct  based on the UPGMA algorithm, with 100 bootstrap replicates to assess branch support:

tree <- aboot(aa.genlight, tree = "upgma", distance = bitwise.dist, sample = 100, showtree = F, cutoff = 50, quiet = T)

            ## color dendrogram by cultivar

cols <- brewer.pal(n = nPop(aa.genlight), name = "Dark2")

plot.phylo(tree, cex = 0.8, font = 1, adj = 0, tip.color =  cols[pop(aa.genlight)])
```

```
nodelabels(tree$node.label, adj = c(1.3, -0.5), frame = "n", cex = 0.5,font = 1, xpd = TRUE)
        ##legend(35,10,c("CA","OR","WA"),cols, border = FALSE, bty = "n")
legend('topleft', legend = c("1B","ER62","RB1","RB2","RB4","SK","Solo"), fill = cols, border = FALSE, bty = "n", cex = 0.8)
axis(side = 1)
title(xlab = "Genetic distance (proportion of loci that are different)")
#*******************
#3.Principle component analysis
aa.pca <- glPca(aa.genlight, nf = 3)
barplot(100*aa.pca$eig/sum(aa.pca$eig), col = heat.colors(50), main="PCA Eigenvalues")
title(ylab="Percent of variance\nexplained", line = 2)
title(xlab="Eigenvalues", line = 1)

# need package ggplot2 for graphic PCA plot
library(ggplot2)
aa.pca.scores <- as.data.frame(aa.pca$scores)
aa.pca.scores$pop <- pop(aa.genlight)
set.seed(9)
p <- ggplot(rubi.pca.scores, aes(x=PC1, y=PC2, colour=pop))
p <- p + geom_point(size=2)
p <- p + stat_ellipse(level = 0.95, size = 1)
p <- p + scale_color_manual(values = cols)
p <- p + geom_hline(yintercept = 0)
p <- p + geom_vline(xintercept = 0)
p <- p + theme_bw()
p
```

- **Variant calling *on subset data (individual cultivar)***

Mpileup: generate variant from BAM file output format in vcf file (document ID# 1275-1281)

BCFtools: to call variant from vcf file Pval threshold 0.5

Filter data on any column using simple expressions (Galaxy Version 1.1.0): to Filter out missing ALT loci (c5==".")

BCFtools: to filter high quality SNP with DP>5, and GQ>20 result high quality SNP/INDEL variants

SnpSift Filter: find common variant genotype (position) among cultivar population: using SnpSift (SnpSift Filter: (countHom() >=) | (countHet() >=)

VCFtool: find specific variant genotype (position) for each cultivar (VCFtool intersect multiple VCF dataset).

SnpSift interval Filter variants **(Snpeff tool in Galaxy):** to select variants in VCF file that locate in gene region (from annotation.cpa.gff Plaza website database) result significant SNP/INDEL variant close to genes.

Merge VCF file and then use R script for visualising data (construct heat map)

```
# DNA finger printing for each cultivar, visualise VCF file by heatmap
# Download Vcf file from Galaxy platform (document# 1425) and check data
mergevcf <- read.vcfR ("Galaxy1425.vcf")
head(mergevcf)

# Extract field GT (Genotype) from Vcf file and check dataset
gt <- extract.gt(mergevcf, element="GT", as.numeric=F)
rownames(gt)
colnames(gt)
head(gt)

# Prepare data for heatmap drawing; convert genotype to numeric
testgt <- gt
testgt[testgt==" "]<- "0"
testgt[testgt=="0/0"]<- "1"
testgt[testgt=="0/1"]<- "2"
testgt[testgt=="1/0"]<- "3"
testgt[testgt=="1/1"]<- "4"
sapply(testgt, mode)

#convert to numeric and make matrix
as.data.frame(lapply(testgt, as.numeric))
mode(testgt) <- "numeric"
testgt_matrix <- data.matrix(testgt)




# Draw heatmap
```

```
testgt_heatmap <- heatmap(testgt, Rowv=NA, Colv=NA, col = colorRampPalette(c("red", "white", "blue"))(256), scale="column", margins=c(5,10))

testgt_heatmap <- heatmap(testgt_matrix, Rowv=NA, Colv=NA, col = colorRampPalette(c("red", "white", "blue"))(256), scale="column", margins=c(5,10))

heatmap.bp(testgt_matrix[100:175,])   #draw only row 100-175


##################################################################
```

Result

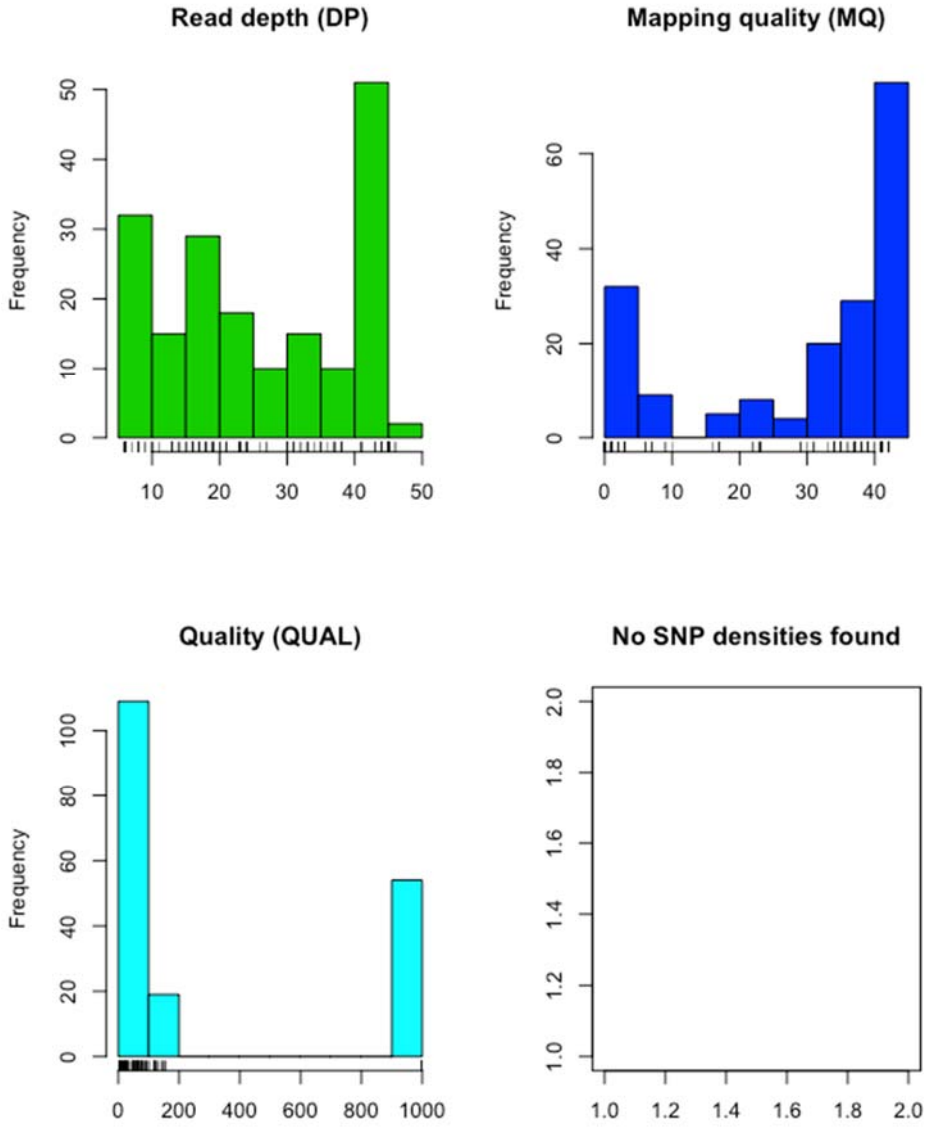

Figure1: Statistic summary for quality control

Table1: Number of variants/polymorphism detected within Australian papaya cultivars

	SNP	INDEL	Total variants
Total number of variants detected	215,402	56,664	272,066
Total number of high quality SNP/INDEL	17,949	796	18,745
Total number of SNP/INDEL located within gene regions	5,605	4,087	9,692

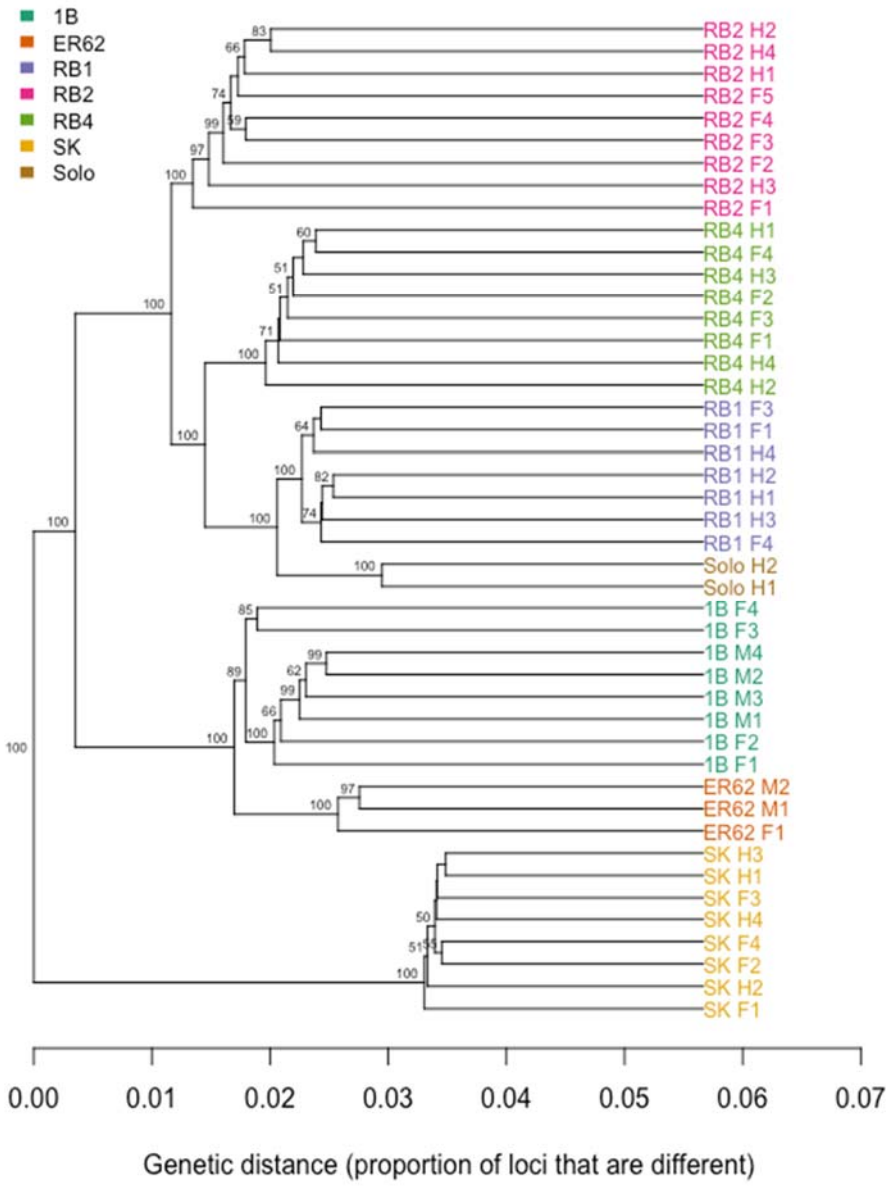

Figure2: Cluster dendrogram of Australian papaya cultivars

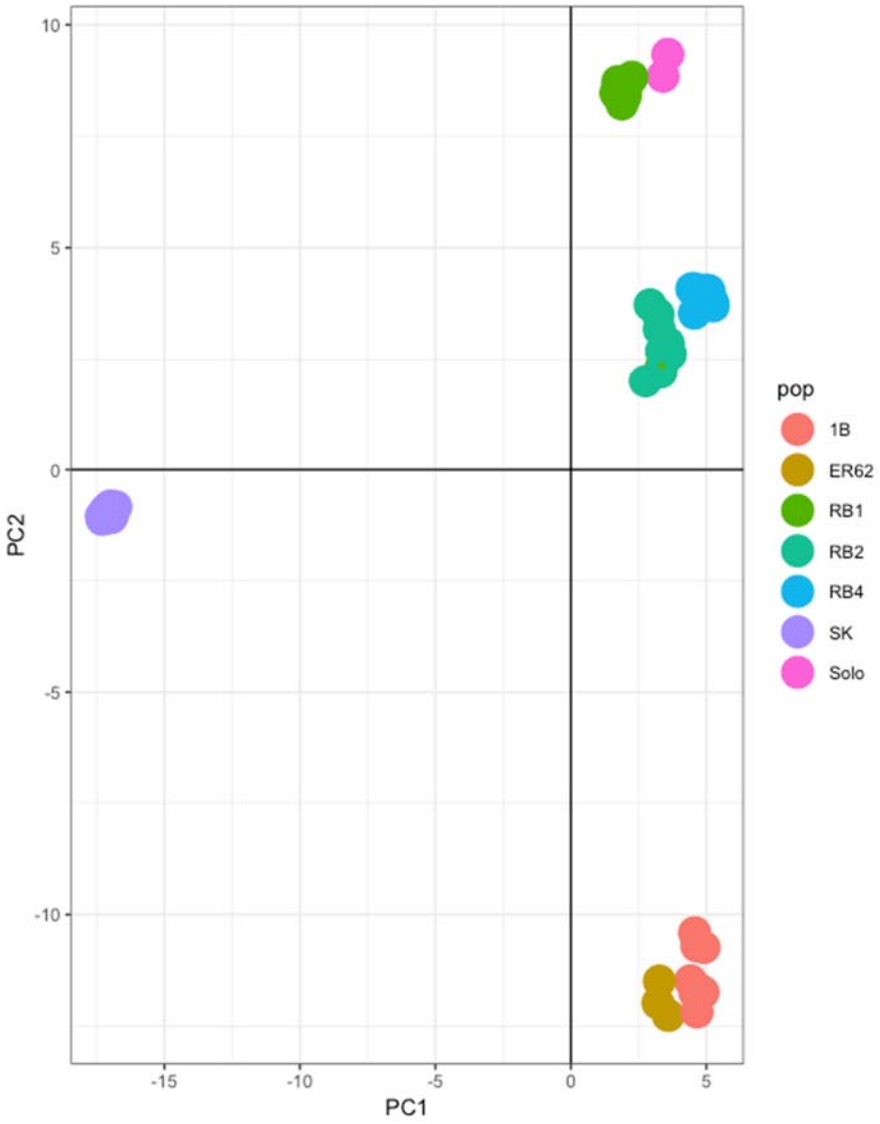

Figure3: Principle component analysis (PCA) of Australian papaya cultivars

Table2: Number of common variants shared among samples of each cultivar

Cultivar	SNP	INDEL	Total variants
1B	272	139	411
ER62	546	81	627
RB1	89	94	183
RB2	158	132	290
RB4	62	110	172
SK	1480	207	1687
Solo	50	154	204

Table3: Number of common variants which located within gene regions

Cultivar	SNP	INDEL	Total variants
1B	130	77	207
ER62	303	39	342
RB1	50	47	97
RB2	84	79	163
RB4	29	70	99
SK	828	112	940
Solo	31	79	110

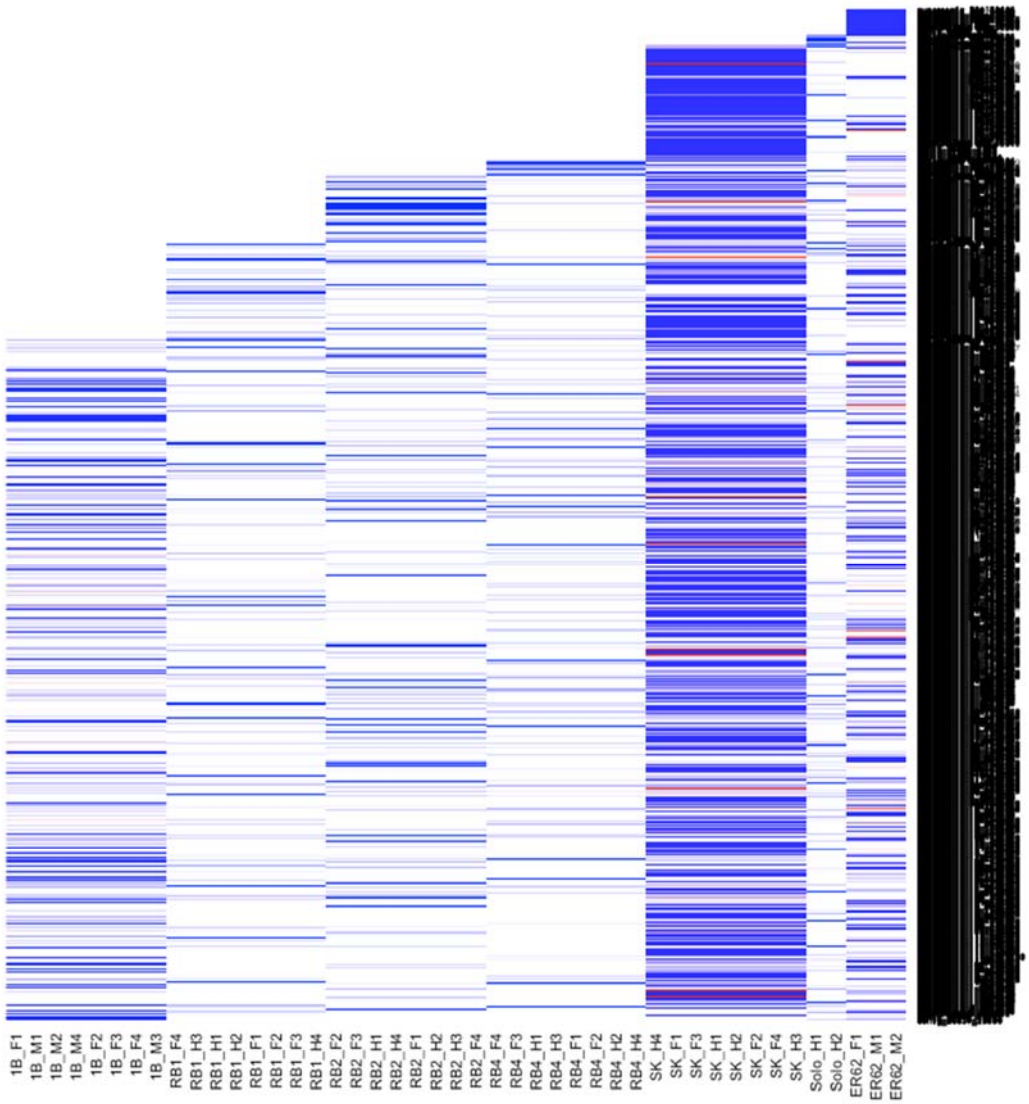

Figure4: DNA finger print of Australian papaya cultivars (heatmap was shown SNP/INDEL loci located within gene regions)

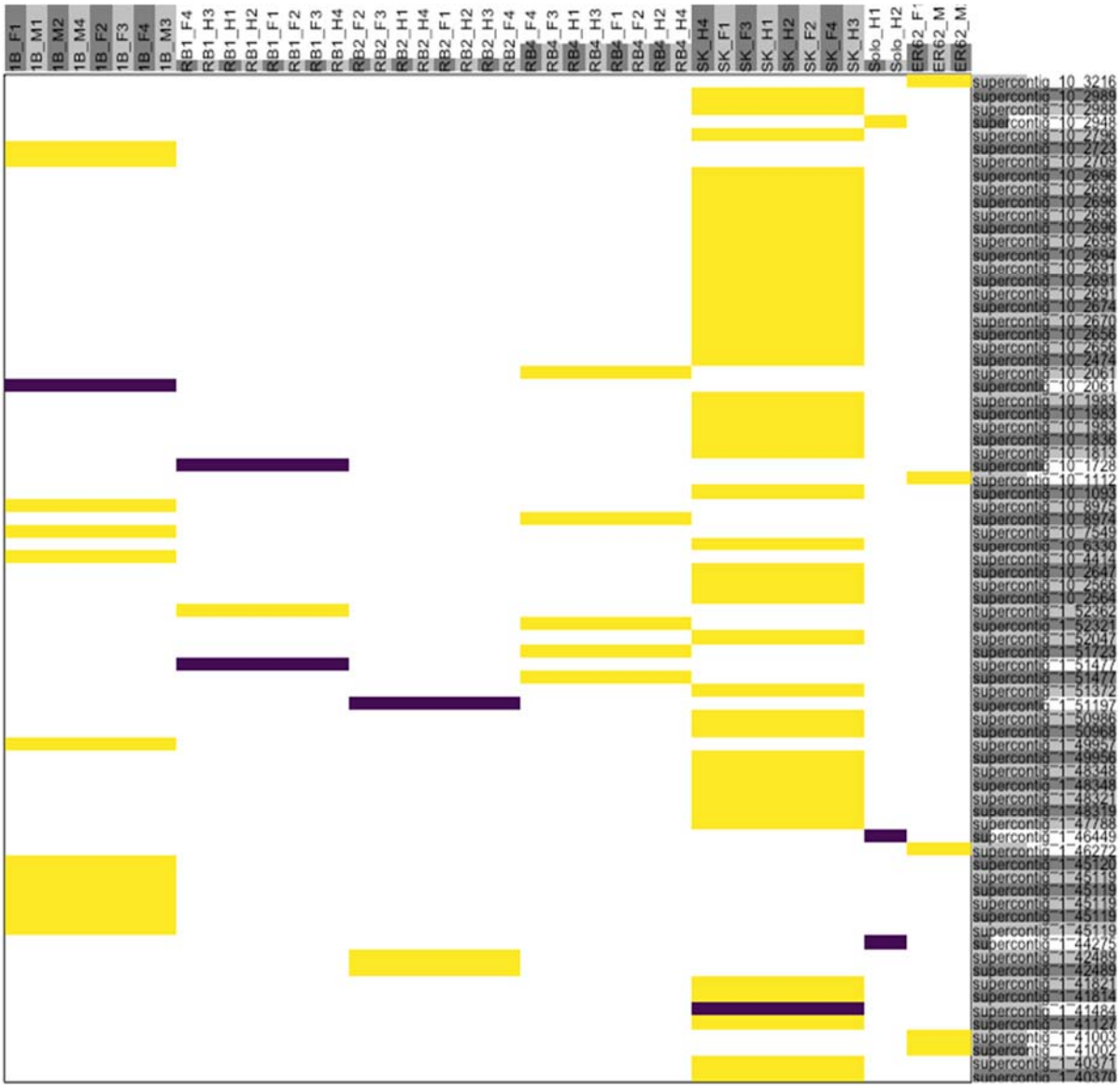

Figure4: DNA finger print specify to each papaya cultivar (heatmap was shown the first 100 SNP/INDEL loci located within gene regions). Yellow colour represent homozygous allele and purple colour represent heterozygous allele.

Reference

Catchen J, Hohenlohe PA, Bassham S, et al. (2013) Stacks: An analysis tool set for population genomics. Mol Ecol 22:3124–3140. doi: 10.1111/mec.12354

Catchen JM, Amores A, Hohenlohe P, et al. (2011) Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. G3 (Bethesda) 1:171–182. doi: 10.1534/g3.111.000240

Ming R, Hou S, Feng Y, et al. (2008) The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). nature 452:991–996. doi: 10.1038/nature06856

Ming R, Yu Q, Moore PH (2012) Papaya Genome and Genomics. In: Genomics of

Peterson BK, Weber JN, Kay EH, et al. (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135. doi: 10.1371/journal.pone.0037135

Ruden DM, Cingolani P, Patel VM, et al. (2012) Using *Drosophila Melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. toxicogenomics 3:35. doi: 10.3389/fgene.2012.00035

Andrews, S.. FastQC A Quality Control tool for High Throughput Sequence Data. [Link]

Li, H. and Handsaker, B. and Wysoker, A. and Fennell, T. and Ruan, J. and Homer, N. and Marth, G. and Abecasis, G. and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. In *Bioinformatics, 25 (16), pp. 2078â 2079.* [doi:10.1093/bioinformatics/btp352][Link]

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. In *Bioinformatics, 27 (21), pp. 2987â 2993.* [doi:10.1093/bioinformatics/btr509][Link]

Cingolani, Pablo and Patel, Viral M. and Coon, Melissa and Nguyen, Tung and Land, Susan J. and Ruden, Douglas M. and Lu, Xiangyi (2012). Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. In *Frontiers in Genetics, 3.* [doi:10.3389/fgene.2012.00035][Link]

Further detail for R can be obtained in Github tutorial (https://galaxyproject.github.io/training-material/topics/sequence-analysis/)

Agreement for the sharing of *Vasconcellea* spp. and *Carica papaya* sequences owned by Griffith University with National Taiwan University

This agreement pertains to the sharing of *Vasconcellea* spp. and *Carica papaya* sequences owned by Griffith University with National Taiwan University within a collaborative project on PRSV-P resistance. The researcher who will receive the sequences is required to agree to the following conditions by signing this form, keeping a copy and sending a copy to Associate Professor Rebecca Ford at Griffith University.

The Taiwan National University understand and agree that:

1. The data sequences supplied by Griffith University of three *C. papaya* (RB2, Solo, and 1B cultivars), *V. pubescens*, *V. parviflora* and *V. quercifolia* genomes remains the intellectual property of Griffith University. Each plant species contain 2 fastq files as details in Table 1.

Table 1: Details of plant species and files names

Species	File names	
C. papaya '1B'	1B_1.fastq	1B_2.fastq
C. papaya 'RB2'	P27-1_1.fastq	P27-1_2.fastq
C. papaya 'Solo'	P30-9_1.fastq	P30-9_2.fastq
V. pubescens	PUB_1.fastq	PUB_2.fastq
V. parviflora	PAR_1.fastq	PAR_2.fastq
V. quercifolia	QUC_1.fastq	QUC_2.fastq

2. All sequences supplied by Griffith University will remain as confidential and will not be distributed further than the immediate research team within National Taiwan University. Sequences supplied will not be submitted to any external databases or published in any way without gaining signed permission from Griffith University.
3. All National Taiwan University staff involved in analysis and writing of any publication stemming from the data will be listed as a co-authorship.

Signed

Shih-Shun Lin

Ford

Associate Professor Shih-Shun Lin
(on behalf of the National Taiwan University team)
Signed Date:

Associate Professor Rebecca Ford
(on behalf of the Griffith University team)
Signed Date: 07/09/2016