Decoding the molecular mechanism of parthenocarpy in *Musa* spp. through protein-protein interaction network

Backiyarani Suthanthiram

National Research Centre for Banana

Sasikala Rajendran

National Research Centre for Banana

Sharmiladevi Simeon

National Research Centre for Banana

Uma Subbaraya (✉ umabinit@yahoo.co.in)

National Research Centre for Banana

Research Article

Keywords: Banana, Parthenocarpy, Candidate genes, Protein-protein interaction network, Pollination

Posted Date: March 16th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-289913/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on July 16th, 2021. See the published version at <https://doi.org/10.1038/s41598-021-93661-3>.

Abstract

Banana, one of the most important staple, delicious fruit among global consumers is highly sterile owing to natural parthenocarpy. Identification of genetic factors responsible for parthenocarpy would facilitate the conventional breeders to improve the seeded accessions. We have constructed Protein-protein interaction (PPI) network through mining differentially expressed genes and the genes used for transgenic studies with respect to parthenocarpy. Based on the topological and pathway enrichment analysis of proteins in PPI network, 12 candidate genes were shortlisted. By exploring the PPI of candidate genes from the putative network, we postulated a putative pathway that bring insights into the significance of cytokinin mediated CLV-WUSHEL signaling pathway in addition to gibberellin mediated auxin signaling pathway in parthenocarpy. Further validation of candidate genes in seeded and seedless accession of *Musa* spp using qRT-PCR put forward *AGL8*, *MADS16*, *IAA (GH3.8)*, *RGA1*, *EXPA1*, *GID1C*, *HK2* and *BAM1* as possible target genes in natural parthenocarpy. In contrary, expression profile of *ACLB-2* and *ZEP* is anticipated to highlight the difference in artificially induced and natural parthenocarpy. Our analysis is the first attempt to identify candidate genes and to hypothesize a putative mechanism that bridges the gaps in understanding natural parthenocarpy through protein-protein interaction network.

Introduction

The term parthenocarpy refers to ovary developing into a seedless fruit in the absence of union of female and male gametes. It has been reviewed in large number of horticultural crops such as grape, tomato, mandarins, banana, opuntia, pepino, eggplant, cucumber and capsicum¹ and stated that parthenocarpy can be achieved as a result of over expression of endogenous hormones in the ovary² and can be genetically controlled^{3,4}. From the inheritance pattern of parthenocarpy in various crops, it has been reported that the trait parthenocarpy is governed by a single dominant gene in eggplant^{5,6}, single recessive gene like in *Capsicum annum*⁷, more than two recessive genes in tomato^{8,9}, a single dominant gene in pepino¹⁰, a single incompatible dominant gene in cucumber^{11,12}, two major additive, dominant-epistatic genes in cucumber¹³, three independent but complementary dominant genes in banana^{14,9}. Phytohormones such as auxin and GA (Gibberellin) playing predominant roles in parthenocarpic fruit development such as tomato¹⁵, *Arabidopsis*^{16,17}, apple¹⁸ etc., It is also being commercially exploited in horticulture crops¹⁹ through exogenous use of irradiated pollen, natural or synthetic hormones such as auxins, gibberellins, IAA etc., during ovary development^{20,21,22}. In spite of so many reports, the molecular mechanism involved in natural parthenocarpic fruit development is still unclear and candidate genes for the trait parthenocarpy have not been identified till date. To understand the molecular mechanism involved in parthenocarpic fruit development, comparative transcriptome analysis has been studied between parthenocarpic and non-parthenocarpic (seeded) accessions in many horticultural crops such as eggplant²³, citrus²⁴, litchi²⁵, oil palm²⁶ etc. Many researchers have tried to identify the parthenocarpic mechanism by studying the expression profile of induced parthenocarpic fruit either via exogenous application or through mutation or genetic transformation^{27,28}.

Among the horticulture crops, banana an economically important crop but its seediness hinders its improvement through conventional breeding approach. Unlike in other crops, ploidy status, intra and inter specific hybridity nature of commercial cultivars/varieties have led to chromosomal imbalance during gamete formation that plays a determinant role in seedless fruit formation. Only limited studies are available for understanding the genetics of parthenocarpy in banana and plantains. It has been stated that the trait parthenocarpy in banana is governed by three independent complementary genes in which the absence of even one dominant gene that resulted in seediness²⁹. Similarly, based on the segregating pattern, it has also substantiated that parthenocarpy is governed by three genes³⁰. Further it has been postulated that among the ancestor genome (A and B) of the present day commercial cultivars, "A" genome coming from *Musa acuminata* (AA) contributes to the female sterility resulting in vegetative parthenocarpy^{31,32,30}. However the loci or the genetic factors responsible for the trait parthenocarpy are not yet identified because of their inherent nature like male and or female sterility, heterozygous nature of parents, unreduced gamete formation etc.

The lack of availability of data associated to seeded and seedless accessions of *Musa* spp hampered perceiving the knowledge on genetic mechanism/factors involved in parthenocarpy. In such scenario, "omics" information related to parthenocarpic trait of various species which are hugely deposited in public databases could be exploited through computational approaches. Several *in-silico* methods such as sequence similarity, evolutionary relationship, detection of SNPs, high throughput gene expression analysis and protein-protein interactions (PPI) etc., could be applied for identifying the genetic factors responsible for parthenocarpy in *Musa* spp. Of which computational prediction of PPI from the gene expression profiles has been widely implemented for the prediction of candidate genes that regulate any complex trait³³. Hence in this study we focused on "proteogenomics" approach by mining the differentially expressed genes (DEGs) of seeded and artificially induced parthenocarpic fruits of various crops, tomato, eggplant, capsicum, grapes, citrus, apple etc., for the identification of candidate genes responsible for parthenocarpy in *Musa* spp. Genetic factors from various orthologous species involved in the parthenocarpic fruit formation and their respective homologous genes in *Musa* spp were taken for the construction of PPI network for the trait parthenocarpy, since it is evidenced that PPIs are conserved in different orthologous species³⁴. The shortlisted genes were validated in seeded and seedless accessions of banana to identify the candidate genes for natural parthenocarpy in banana.

Results

Construction of PA-PPI network

A total of 210 DEGs were extracted from the transcriptome profile of various crops with respect to artificially induced parthenocarpy and from the genetically modified crops. BLAST analysis of these genes against *Musa* ssp (<https://banana-genome-hub.southgreen.fr/blast>) displayed a hit with 156 orthologous protein sequences with an identity of $\geq 70\%$ [Supplementary Material 1]. An initial PPI network was constructed for these 156 orthologous protein sequences that resulted in the formation of a putative network with 95 nodes and 185 edges. Further, as a result of Agilent literature search (63 proteins) an additional 49 nodes with 34 edges were merged to the initial network [Supplementary Material 1] to get an extended PA-PPI network of 140 nodes (Proteins) with 219 edges (interactions) and designated as undirected network [Supplementary Fig. S1]. Structural properties of the constructed network such as number of nodes, degree distribution, clustering coefficient were calculated [Supplementary Table S1] for better understanding the functional organization of proteins in the network. For example, the average connected component in PA-PPI network was found to be 22, indicating that the majority of the proteins in the network are highly connected that play a central role in the network's architecture and considered to be essential proteins³⁵. Further the degree distribution in this network approximates the power law ($P(k) \sim (k - \gamma)$) ($P(k) \sim (59.5 - 1.3)$) i.e., with smaller value of degree exponential (γ) -1.3. This in turn determined the importance of hubs in the network i.e. networks with larger γ (< 3) value indicate that the hubs in the network are relevant to biological function rather behave like a random network. This meant that our network could perfectly reflect a biological network and thereby proteins in the network might efficiently communicate biological information related to parthenocarpy. In addition, the PA-PPI network has a characteristics average path length value of 5 and comprised 40% shortest paths. This outlined the overall navigability of the network that the biological information in the network could get transferred by crossing few nodes from a selected protein to others in the network³⁶. The clustering co-efficient of this scale free network is 0.283, that significantly describes that the internal structure of this network is highly interactive and form clusters.

Topological Analysis Of Pa-ppi

Biological significance of proteins in this scale free network was determined by analyzing the centrality measures (topological properties) such as degree, betweenness and closeness centrality. Based on the topological properties of the PA-PPI network, top ten proteins with higher degree, higher betweenness centrality scores and higher closeness centrality score were taken and listed in Table 1. The average degree of proteins in the PA-PPI network was found to be 3.128 and proteins with high degree (> 10 interacting partners) such as *LFY*, *ZEP*, *HK2*, *EXPA1* and *SL1* are referred as degree based hubs. Proteins with higher betweenness centrality scores such as *NIA1*, *ZEP*, *FL*, *NCED1*, *MOCOS* could act as useful indicators for detecting bottleneck protein in the PA-PPI network. *LFY*, *FIE2*, *GAF1*, *NFYB9*, *ZEP* with high closeness centrality have a smaller path length to reach all other proteins in the network and thereby these proteins would have a greater influence in the network. In the overall topological analysis, it was found that some of the genes such as *LFY*, *ZEP*, *EXPA1*, *SL1*, *GH3.8*, *BAM1*, *HK2* etc., were repeatedly present in all the three centrality measures and thus these genes were shortlisted as potential candidate genes.

Table 1
Topological Analysis- Degree, Betweenness, Closeness centrality and Cluster Analysis

A) Ranked by Degree			B) Ranked by Betweenness method			C) Ranked by Closeness method			D) Ranked by MCC method		
Rank	Node	Score	Rank	Node	Score	Rank	Node	Score	Rank	Node	Score
1	<i>LFY</i>	21	1	<i>NIA1</i>	3936	1	<i>LFY</i>	39.2159	1	<i>ZEP</i>	725762
2	<i>ZEP</i>	12	2	<i>ZEP</i>	3694.22	2	<i>FIE2</i>	32.406	1	<i>GAF1</i>	725762
2	<i>GAF1</i>	12	3	<i>LFY</i>	3602.76	3	<i>GAF1</i>	31.9071	3	<i>EXPA1</i>	725761
4	<i>EXPA1</i>	11	4	<i>NCED1</i>	3534	4	<i>NFYB9</i>	31.0825	3	<i>RAP23</i>	725761
4	<i>RAP3</i>	11	5	<i>MOCOS</i>	3480	5	<i>ZEP</i>	30.1262	5	<i>HK2</i>	725760
6	<i>HK2</i>	10	6	<i>GAF1</i>	3319.09	6	<i>MADS2</i>	29.5825	5	<i>At4g13710</i>	725760
6	<i>At4g13710</i>	10	7	<i>FIE2</i>	2807.91	7	<i>EMF2</i>	29.4159	5	<i>BAM1</i>	725760
6	<i>BAM1</i>	10	8	<i>E2FB</i>	1994.8	7	<i>MAD16</i>	29.4159	5	<i>SL1</i>	725760
6	<i>SL1</i>	10	9	<i>PHSH</i>	1609.17	9	<i>EXPA1</i>	29.1833	5	<i>GH3.8</i>	725760
6	<i>GH3.8</i>	10	10	<i>ANT</i>	734.21	10	<i>AP2</i>	28.5762	10	<i>SCL7</i>	362880

Cluster Analysis

Highly interconnected regions or sub network in PA-PPI were identified using MCODE plug-in since clusters in a network are often protein complexes which involved in the same pathway and the same protein family. Totally eight clusters were obtained of which only 3 clusters having a score of > 3 were subjected to functional enrichment analysis using ShinyGO (Fig. 1 & [Supplementary Fig. S2A]. Cluster 1 comprises of genes belonging to the biological process such as "Response to hormone-mediated pathway, Response to hormonal/chemical stimulus, Regulation of multicellular organismal development and Regulation of gene expression" etc., [Supplementary Fig. S2B]. Cluster 2 encompasses genes belonging to embryo, seed sac development, mitotic cell cycle, Gametophyte development, reproductive processes etc. Interestingly genes belongs to cluster 3 involved in Carbohydrate metabolic process, Oxido-reduction coenzyme metabolic process, Energy reserve metabolic process etc., [Supplementary Fig. S2C]. The unique rankings

of genes based on each centrality measures and MCODE clusters analysis are given in Table 1 and the genes belongs to each clusters as a result of MCODE plugin are given in [**Supplementary TableS2**].

Validation Of Candidate Genes For Parthenocarpy

Comprehensive analysis of the network highlighted various hubs and essential genes that could directly or indirectly associated with parthenocarpy were mined. Based on the following observations, nodes from the PPI network were taken as candidate genes

1. Nodes which are repeatedly ranked as top nodes in each centrality measures
2. Nodes that are listed in top clusters from MCODE analysis
3. Nodes that are commonly found in both topological and cluster analysis

The shortlisted genes namely *ZEP*, *LFY*, *MADS29*, *GID1C*, *RGA1*, *HK2*, *MADS16*, *BAM1*, *GH3.8*, *AGL8*, *EXPA1*, and *ACLB2* were subjected to experimental validation using qRT-PCR. (**Table 2**). Ovary samples of seeded *Musa acuminata* ssp. *burmanicoides* Calcutta 4 - C4) and seedless (Pisang Lilin -PL; cv. Rose - CVR) accessions of *Musa* spp under three different conditions namely un-pollinated (UnP), 24 hrs (P24) and 48 hrs (P48) after pollination (Fig. 2) were taken for the present study. Under un-pollinated condition, *MADS16*, *GH3.8*, *LFY*, *EXPA1* and *RGA1* exhibited similar expression pattern in CVR and PL compared to C4 (Fig. 3). Whereas genes such as *HK2*, *BAM1* and *MADS29* exhibited similar pattern of expression in C4 and CVR compared to PL (Fig. 3). Whereas *AGL8*, *GID1C* and *ACLB2* genes exhibited same pattern in all the three cultivars under un-pollinated condition. Under pollination after 24 hrs (P24), genes namely *AGL8*, *MADS16*, *RGA1*, *BAM1*, *ACLB2*, *EXPA1* and *LFY* exhibited similar level of expression in C4 and CVR compared to PL. *MADS29*, *HK2* and *IAA* (*GH3.8*) had similar level of expression in CVR and PL compared to C4 (Fig. 3). It is interesting to note that unlike other genes, expression of *GID1C* was very specific to CVR. Further at P48 hrs conditions, *GID1C*, *MADS29*, *ACLB2*, expansin exhibited similar expression in C4 and CVR. Other genes *AGL8*, *MADS16*, *RGA* (*DELLA*), *BAM1*, *HK2* and *LFY* exhibited distinct expressions in C4, CVR and PL with no pattern under P48 (Fig. 3). Among the 12 genes, Ct value of *ZEP* is “undetermined” in all the three cultivars under both pollinated and un-pollinated conditions. Overall, it was found that genes such as *AGL8*, *GID1C*, *MADS16*, *MADS29*, *BAM1*, *HK2*, *IAA* (*GH3.8*), *RGA1* had similar expression pattern in *Musa* spp as reported in other artificially induced parthenocarpic horticultural crops such as tomato³⁷, eggplant³⁸, pear³⁹, apple⁴⁰.

Discussion

Identification of candidate gene/s responsible for the trait parthenocarpy in *Musa* spp. is a prerequisite for the commercial exploitation of accessions with B genome that exhibit resistance against biotic or abiotic stresses⁴¹. It was reported that the trait parthenocarpy in banana is governed by three independent complementary genes where seediness could result even with the loss of one dominant gene^{29,30}. But till date, only limited research has been attempted in banana to understand the genetic basis of natural parthenocarpy. Further candidate genes documented from various studies and crops are related to artificially induced parthenocarpy but not on natural parthenocarpy as in banana. Thus computational method of protein-protein interaction networks based data mining was employed for the identification of candidate genes in banana^{42,43,44,45&46}. Since most of the biological activities in the cell are result of molecular interaction/s among bio-molecules and thus study of PPI is essential to unravel the molecular basis of any complex traits⁴⁷. Differentially expressed genes and genes with respect to parthenocarpy from various crops were mined and their respective orthologous genes in *Musa* spp were retrieved for the construction of parthenocarpy associated PPI in consequent to a study that orthologous sequences are ought to have the same functions⁴⁸. [**Supplementary Fig. S1**].

Table.2. Shortlisted candidate genes from PA-PPI network for the trait parthenocarpy

Accession ID	Gene Name	Gene Description	Expression of gene in other seedless traits**	Expression in <i>Musa</i> spp.								
				C4			CVR			PL		
				UnP	P24	P48	UnP	P24	P48	UnP	P24	P48
GSMUA_Achr7P18880_001	<i>ZEP</i>	Zeaxanthin epoxidase	Down regulation	*	*	*	*	*	*	*	*	*
GSMUA_Achr6P16390_001	<i>LFY</i>	LFY-like protein OrcLFY	Down Regulation	++	—	—	-	—	-	-	-	-
GSMUA_Achr3P23580_001	<i>MADS29</i>	MADS-box protein AeAP3-2 isoform X1	Down Regulation	-	++++	++	—	+++	—	++	+++	+
GSMUA_Achr3P22920_001	<i>HK2</i>	probable histidine kinase 2	Down Regulation	++	+	+	+	—	—	—	—	—
GSMUA_Achr9P20950_001	<i>MADS16</i>	MADS-box transcription factor 16-like	Down Regulation	++	++	-	-	++	++	-	—	-
GSMUA_Achr4P07370_001	<i>BAM1</i>	leucine-rich repeat receptor-like serine/threonine-protein kinase BAM1	Down regulation	++	-	-	++	-	-	-	—	—
GSMUA_Achr4P07220_001	<i>GH3.8</i>	probable indole-3-acetic acid-amidosynthetase GH3.8	Down regulation	++++	++	++	+	—	++	+	—	++
GSMUA_Achr11P05030_001	<i>ACLB-2</i>	ATP Citrate synthase beta chain	UP regulation	-	++	++++	-	+++	+++	-	—	-
GSMUA_Achr3P02280_001	<i>AGL8</i>	Agamous like MADS TF	Down regulation	++	++	-	++	+++	++++	++	—	++
GSMUA_Achr1P21300_001	<i>RGA1</i>	DELLA	Down regulation	+++	++	+	++	+	—	+	—	+
GSMUA_Achr8P05910_001	<i>GID1C</i>	Gibberellin 1C receptor like	UP regulation	+++	++	-	+++	-	-	++++	+++	++
GSMUA_Achr1P02650_001	<i>EXPA1</i>	expansin-A11-like	UP regulation	+++	-	++	+	—	—	+	+++	+

** References included in Supplementary files

* qRT-PCR expression values is “Undetermined”

+ → Expression; ++, +++ & ++++ → 2, 3 and 4 fold expression

- →Down regulation; -, — & — →2, 3 and 4 fold down regulation

Functional enrichment of the proteins in the overall network revealed that majority of the genes are primarily involved in floral whorl development (26.32%), meristem maintenance (15.79%), regulation of reproductive process (10.53%), transcriptional regulation, gene regulation (10.52%) and oligosaccharide biosynthetic process (10.53%). This in turn supported the relevance of orthologous genes short listed for the construction of PA-PPI network in the current study since many studies reported the significance of genes involved in floral development, ovule integument, reproductive process etc., in seedless fruit formation^{49,50}. Considering the functional and GO analysis of PA-PPI network together, it was shown that majority of the genes that framed the PA-PPI are involved in “regulation of cellular macromolecule biosynthetic process” and “transcriptional regulatory activity” [Supplementary Fig. S3]. Particularly transcription factors that are widely involved in the flower - fruit transition stage like *AGL8*, *MADS16*, *LFY* and *MADS29* get highlighted as a result of centrality measures. Similarly KEGG pathway analysis rationalized that hormone signal transduction, carotenoid biosynthesis, fatty acid metabolism, carbohydrate metabolism and lysine degradation pathways having a strong association in the network of natural parthenocarpy (Fig. 4). In our previous review, the role of hormone mediated transcriptional regulation in parthenocarpy was emphasized⁶¹, while the current study highlighted involvement of proteins in carbohydrate, fatty acid and lysine degradation pathways, however their interrelation in inducing parthenocarpy remains elusive [Supplementary Table S3]. It has been speculated that some of the discrete nature of parthenocarpic fruits such as its nutritional value, pulp content, fruit size, peel thinness etc., might be due to the cellular metabolism that occurs in parthenocarpic fruit formation.

While considering the role of lysine degradation pathway, it was found that glycine and carnitine are the end products which incite us to acquire information regarding free amino acid (FAA) content difference in parthenocarp and seeded varieties of *Musa* spp. Variation in the level of FAA between parthenocarp and seeded traits were analyzed in

tomato and capsicum but still the possible role of FAA content in parthenocarpic fruit formation is yet to prove. On the other hand, genes namely *MEA*, *FIE*, *CLF* and *SWN* are group of polycomb (PcG) proteins are also get highlighted under lysine degradation pathway and their interrelation in inducing parthenocarp remains elusive.

PcG proteins are act as histone modifying enzymes and reported to regulate the embryo and endosperm proliferation and anteroposterior organization during seed development⁵¹. Functional mutation of *FIE*⁵² *FIS*⁵³ and *MEA*⁵⁴ were also observed in parthenocarpic fruit development⁵⁵. In addition interaction of MADS box TFs namely *AP2*, *AGL15*, *AGL2/EMF*, with PcG proteins namely *MSI1*, *FIE*, *SWN* and *VEL1* through *VRN5* (Vernalization 5) and *TPL* (transducing family protein/WD-40 repeat family protein) was found in the putative network. Down regulation of *VRN5*, *TPL* as well MADS TFs and PcG proteins were already reported in parthenocarpic fruit development⁵⁶ suggesting their promising role. Despite these reports, the direct role of polycomb (PcG) proteins in lysine degradation pathway remains unclear in seed development and understanding the integration of these genes and the pathways in parthenocarp is the key challenge. Another unpublished work at ICAR-NRCB, reported failure of certain female fertile accessions to set seeds under a set of environmental condition but the reason behind this behavior remains undiscovered. Thus it is speculated that PcG might be involved in epigenetic mechanism that regulates the seed formation under specific environmental conditions.

Candidate genes mined as a result of network topological analysis were validated through expression analysis and their interacting partners were explored using Bisogenet app in Cytoscape in order to fathom their mechanism in parthenocarp. Some studies reported that parthenocarp trait is governed by "A" genome^{29,31}, while in another study the dynamic nature of AA genome accessions exhibiting both seeded (C4) and seedless traits (PL)^{57,58} has also been reported. Among the seedless AA diploid accessions, some are amenable (cv. Matti, cv. Rose) and few are recalcitrant (PL) to seed set upon artificial pollination⁵⁹. Thus in the current study, expression pattern of the candidate genes were compared among three 'AA' accessions namely i) Calcutta 4 (C4) - with profuse seed set ii) cv. Rose (CVR) - rarely setting seeds upon pollination iii) Pisang Lilin (PL) - seldom setting seeds. Ovary samples from the test accessions were collected at three different time intervals, un-pollinated (UnP), 24hrs after pollination (P24) and 48 hrs after pollination (P48) and subjected to real time expression analysis. The expression study revealed that irrespective of the cultivars, majority of the variations was observed between the seeded and seedless accessions at 24 hrs after pollination (P24). This confirmed that ovary sampling around 24 hrs after pollination is optimum for seed set studies in *Musa* spp which could be well correlated to our earlier findings that male gametes reach the ovule within 24 hrs after artificial pollination⁶⁰. Genes belonging to MADS family namely *MADS16*, *MADS29*, *AGL8*, *LFY* were repeatedly found in all the three centrality measures as well as in cluster analysis and thus ranked as key genes. Expression level of the two MADS box transcription factors namely *AGL8* and *MADS16* were up regulated in C4 and CVR and down regulated in seedless PL at P24 condition. This confirms the negative regulation of MADS box transcription factors upon pollination in natural parthenocarpic accessions like PL. This is in correlation to our previous review that unveiled that upon fertilization, MADS box transcription factors in parthenocarpic accessions could act as a key regulator for the transition between the state of "ovary arrest" to fertilization triggered fruit set that mediate seedless fruit formation⁶¹. Next to the MADS box transcription factors, proteins involved in hormone regulatory mechanism (Histidine kinase (*HK2*), indole-3-acetic acid (IAA)-amidosynthetase GH3.8, DELLA (*RGA*, *RGL1*, *RGL2*), Gibberellin receptor (*GID1C*), Expansin (*EXPA1*) and cellular metabolism (ATP citrate synthase (*ACSB2*), zeaxanthin epoxidase (*ZEP*), leucine-rich repeat receptor-like serine/threonine-protein kinase (*BAM1*) scored next top ranks.

The expression level of *RGA* (*DELLA*) in PL was down-regulated by 3 fold compared to C4 and CVR at P24. Negative regulation of *DELLA* in seedless fruit formation is well documented during artificially induced seedless fruit formation in various crops^{40,62,63,64,65,66,67,68}. Another study proved that silencing or mutation of *DELLA* prevents its interaction with other proteins involved in seed development and thereby inducing facultative parthenocarp in tomato⁶⁹. Interactions of proteins associated with *DELLA* (*RGA*) during floral or fruit initiation after fertilization and upon increased GA signaling are reported in several studies related to parthenocarp¹⁸. It is also proved that external application of GA competes *DELLA* for its interaction with *GID1C* leading to subsequent degradation of *DELLA* eventually resulting in seedless fruit formation⁷⁰. Further it is interesting to note that, upon pollination, there was reduction in the expression level of *GID1C*, where the reduction was drastic in CVR followed by C4 and PL⁶⁹ proved the interaction of *GID1C* with *DELLA* which is essential for seed development by studying the various mutants of *GID1C* and *DELLA* in *Arabidopsis*. When associating the expression profile of these four genes namely *AGL8*, *MADS16*, *DELLA* and *GID1C* together with their interacting partners from the network, strongly supported the proposed hypothesis that MADS box TFs together with DELLA could act as a focal point in bridging auxin and GA hormone during seed set. Down regulation of *IAA* (GH3.8) was observed in the ovaries of both CVR and PL at P24 which reiterated its expression in hormone induced parthenocarp. Down-regulated expression of *AUX/IAA* and homologs of *GH3.2* was also reported in parthenocarpic eggplant over the seeded eggplant³³. *IAA* (*GH3.2*) is an auxin-amino acid conjugating enzyme that converts auxin into an inactivate form, reported that homolog of *GH3.2*, (i.e. IAA-amino synthetase (*GH3.8*)) down regulates auxin signaling by preventing the accumulation of free *IAA*^{71,72}. Recent study on *DELLA* and its interaction with *ARF7/IAA9*, shown that *GH3.2* is involved in the regulation of auxin homeostasis through GA during fruit development¹⁸. This supports our earlier review that increased GA together with decreased expression of auxin responsive genes like *GH3.2* resulting in seedless fruit formation^{61,73,74}. GA in coordination with *IAA* was reported to increase the expression of *EXPA1* and thereby promoted cell division and cell expansion^{74,75}. Expansin (*EXPA1*), a growth promoting gene involved in cell division and cell growth, subsequently required for fruit development. GA induced parthenocarp showed increased expression of expansin in the ovaries of pear fruit suggesting that genes involved in cell expansion, cell division get activated upon hormonal signaling for fruit²⁷. Similarly *EXPA1* is up regulated in PL whereas it is down regulated in both C4 and CVR at P24. This

implies that upon pollination, *EXPA1* could positively mediate natural parthenocarpy in *Musa* spp. Histidine Kinase *CK1* (*HK2*), a cytokinin receptor reported to be involved in the biological process of cytokinin signaling pathway, flower development, seed germination, meristem development and in several stress related mechanism⁷⁶. In an earlier report based on comparative genomic approaches, an orthologous gene of *HK2* involved in gametophyte development was reported in the genome of *Musa*⁷⁷. Association of *HK2* in parthenocarpy is not elusive, however the constructed PA-PPI network proposed that *HK2* interacts with proteins associated to seed development (*BAM1*) and cell expansion (*SCL*, *EXPA1*). While another candidate gene, BARELY ANY MERISTEM1 (*BAM1*) received much attention while probing the interaction of *HK2* in the network. Importance of *BAM1* and *BAM2* in male and female gametophyte functionality in *Arabidopsis* further emphasized its importance in seedless fruit formation⁵⁸. Expression study of histidine kinase (*HK2*) and *BAM1* shown that in seedless cultivars (PL and CVR) they were down regulated and the level of expression was much lower in PL compared to CVR. Similar expression profile of these two genes in other seedless fruits such as tomato, eggplant and capsicum further evidenced its role in parthenocarpy⁷⁸. A Genome wide association study for seedlessness in *Musa* sp highlighted that an orthologous gene to Histidine Kinase (*CK1* or *HK*) is strongly related to female sterility and speculated that SNPs found in this gene might be responsible for parthenocarpy⁷⁷. Our study supported that apart from SNPs, the expression of the gene is also playing a major role in parthenocarpy. Probing the interacting partners of *HK2* in the network shown that protein involved in pollen tube elongation (*PI4KB1*)⁷⁹, brassinosteroid (BR) integrated seed development (*NAC081*)⁸⁰, seed storage (*AT2S3*), microtubule-associated protein involved in cell expansion (*DRP1A*)⁸¹ get interacts with *HK2* which in turn clearly evidenced its role in seed development. Further extending our search for experimental interactions linked with another candidate gene, *BAM1* apparently shown *NAP5*, *CLV3*, *CLV1* and *CPI1* as interacting partners. Among the interacting partners, *CLV3* and *CLV1* have gained considerable attention since significant down regulation of *BAM1/2* and WUSCHEL (*WUS*) was observed in the sexual sterility (*Sises*) mutant seedless tomato, exhibiting both male and female sterility⁵⁶ where *CLV* succeeded its interaction in the *CLV*-WUSCHEL signaling pathway that plays a multifunctional role in plant development⁸². WUSCHEL, a transcription factor that mediates the expression of *CLV3* and *AG* during floral development particularly in the ovule and induces integument formation^{83,84,85}. Considering everything, it was suggested that interaction of *BAM1* with *CLV1* mediates *CLV*-WUSCHEL signaling pathway, in which *CLV* activates WUSCHEL which results in the transcriptional activation of MADS box transcription factor, *AG*, leading to seed formation. Further, expression of *MADS29* was drastically increased upon pollination irrespective of the cultivars however its expression is more in the seeded accession C4, this might be due to the importance of *MADS29* in fruit set upon pollination. Increased expression of *ACSB2* was evidenced in artificially induced seedless tomato⁷⁴, however *ACSB2* was down regulated in PL and up regulated in C4 and CVR upon pollination at P24. *LFY*, another candidate gene reported to involve in floral meristem initiation⁸⁶ was observed to down regulated in all the three cultivars upon pollination (both P24 and P48). This inferred that *LFY* might plays a role in floral initiation rather than fruit and seed set. In addition, Ct value of the gene *ZEP* was undetermined due to its in all three accessions irrespective of the conditions. These results from banana interrogated the significance of *LFY*, *ACSB2* and *ZEP* in natural parthenocarpy. Overall, our results emphasized that genes *AGL8*, *MADS16*, *RGA1*, *GIDC1*, *IAA (GH3.8)*, *EXPA1*, *HK2*, and *BAM1* could play significant role in both induced and natural parthenocarpy. Based on the illustration of possible interactions associated with candidate genes in the PA-PPI network together with their expression study (Fig. 3), we designed a putative pathway that is speculated to be the underlying parthenocarpy mechanism in *Musa* (Fig. 5).

Materials And Methods

The approach used in this study for prioritizing key genes in parthenocarpy is summarized in Fig. 6 and described in the following sections.

Mining Of Genes Associated With The Trait Parthenocarpy

Genes associated with parthenocarpy in other crops were mined from databases like Uniprot, KEGG and sources like Pubmed, Pubmed Central, etc. This is achieved through manual text mining by using the query words "parthenocarpy", "seedlessness", "parthenocarpy and genes", "parthenocarpy and transcription factors", "parthenocarpy and *Musa*". In addition, highly enriched differentially expressed genes (DEGs) between seeded and artificially induced parthenocarpic fruits either through phyto-hormone or chemical spray/mutation/genetic transformation in various crops such as tomato⁸⁷, eggplant³³, apple²⁸, pear²⁸ etc., were extracted from their respective transcriptome profiles [Supplementary Material 1].

Retrieval of orthologous sequences in *Musa* spp

The corresponding sequences pertaining to the mined parthenocarpic genes were downloaded in fasta format either from Uniprot (<https://www.uniprot.org/>) or from the respective crop specific genome or transcriptome databases using their unique reference gene ID cited in the literature. These sequences were then submitted to BLAST search in Banana Genome hub (<http://banana-genome-hub.southgreen.fr/>)⁸⁸ in order to retrieve corresponding orthologous sequences in *Musa* species which has $\geq 70\%$ sequence identity [Supplementary Material 1].

Construction Of Protein-protein Interaction Network (ppi)

The retrieved *Musa* orthologous sequences were submitted to STRING v10.5 (<https://string-db.org/>), a pre-computed database for the exploration of PPI⁸⁹. Predicted protein association networks with a combined score of > 0.4 were taken for the construction of PPI network using Cytoscape 3.7.1⁹⁰. Since the initial PPI network constructed using STRING database had limited number of nodes (proteins) and edges (interactions) for the study, we extended our search of possible interacting partners for the extracted genes using the plugin called Agilent Literature Search⁹¹ in Cytoscape [Supplementary Material 1].

Topological and cluster analysis of the network

The extended PPI network is considered as an undirected graph (G) constituting the components V and E, in which proteins are denoted as nodes (V) and the interactions are represented as edge (E). In the current study, to identify key proteins from the network, topological properties such as degree (k), betweenness centrality (BC) and closeness centrality (CC) were analysed. These three different centrality measures were calculated using CytoHubba, a Cytoscape plugin (Chin et al. 2014) that explored nodes with high degree, high BC and CC to identify the important proteins related to parthenocarpy from PPI network. Cluster analysis was performed using Molecular Complex Detection (MCODE) plug-in which provides a novel clustering algorithm to screen the modules of the PPI network for parthenocarpy (PA-PPI) through Cytoscape (Bandettini et al. 2012). MCODE scores of > 3 and the number of nodes > 3 were set as cutoff criteria with the default parameters (Degree cutoff ≥ 2 , Node score cutoff ≥ 2 , K-core ≥ 2 and Max depth = 100). Genes identified from the clusters and the top ten genes from topological analysis were subjected to Gene ontology (GO) and KEGG pathway enrichment analyses using BinGO, ClueGO and BLAST2GO in order to expedite the functional annotation of each genes ⁹².

Identification And Collection Of Plant Materials

The test samples were collected from the field Musa genebank of ICAR-National Research Centre for Banana (NRCB), Tiruchirapalli, Tamil Nadu, India where more than 300 Indian accessions and 121 exotic accessions are being maintained. The tissue cultured propagules (AA genomic group) of the seeded accession (Calcutta 4 (C4)) and parthenocarpic accessions (Pisang Lilin (PL); and cv. Rose (CVR)) were received from the International Transit Centre (ITC), Belgium through ICAR- National Bureau of Plant Genetic Resources (NBPGR). Exotic collections (EC) numbers were given to the exotic introductions by ICAR-NBPGR and the details for the three cultivars used in the current study are provided below in the Table 3.

Table 3

Accession number, genomic and parthenocarpic nature of the accessions used in the study.

Accession Name	Accession Number	ITC number	Genomic group	Nature of the group
Calcutta 4	0654	ITC 0249	AA	Seeded
Pisang Lilin	0195	ITC 1121	AA	Parthenocarpy
cv. Rose	0638	ITC 0712	AA	Parthenocarpy
Matti	0182	—	AA	Parthenocarpy

As a standard protocol, this has been deposited with ICAR-NBPGR for *in-vitro* maintenance. Tissue culture plants of these test accessions were sub cultured in the rooting media and the rooted plantlets were acclimatized through primary and secondary hardening under green house. The secondary hardened plants were planted in five replications with five plants per replication in the ICAR- NRCB field and maintained with standard package of practices. At flowering time, the whole inflorescence was bagged before opening of the first female hand. The ovary samples of each accession were collected from the female floret of C4, CVR and PL on the day of flower opening at 8.00am and designated as un-pollinated (UnP) sample. For pollen grains (Male), Matti (AA) a local landrace collected from Thriunelveli, Tamil Nadu, India is being maintained at ICAR-NRCB with accession number 0182. The pollen grains were collected during anthesis at 7.00 am from the accession Matti (Male), dusted over the stigma of the female florets (C4, CVR and PL) and the whole inflorescence was covered. The ovary samples were collected from the female florets at 24 hrs and 48 hrs after pollination and designated as P24 and P48 samples. The collected ovary samples were immediately snap frozen and stored at -800C. ICAR-National Research Centre for Banana, Tiruchirapalli, Tamil Nadu, India being the National Active Germplasm Site (NAGS), all national and international guidelines and legislations were followed in performing the experimental research, field studies and collection of experimental samples.

Validation Of Genes Using Qrt-pcr

From the result of topological and cluster analysis of parthenocarpic PPI (PA-PPI) network, we selected 12 candidate genes and subjected to qRT-PCR to compare their relative expression in seeded and parthenocarpic accessions of *Musa* spp. Total RNA of each sample was extracted using Agilent Plant RNA isolation mini kit (Agilent Technologies, Inc., USA) (product no. 5188–2780) and the quantity /integrity of the RNA was checked using nanodrop (ColibriMicrovolume Spectrometer -Titertek/Berthold). The cDNAs of each RNA sample was synthesized using Transcriptor First strand cDNA synthesis kit (Roche) according to the manufacturer’s instructions. Primers for the experiments were designed using IDT-Primer design tool (<https://www.idtdna.com/pages/tools/primerquest>) and the primer sequences are provided [Supplementary tableS4]. The qRT-PCR reactions were performed in triplicates with Ribosomal protein S2 (RPS2) as endogenous control and repeated thrice on three biological replicates and run on Light-Cycler 96 instrument (Roche Co. Germany) using the SYBR Green Master Mix (Thermo Fisher Scientific, USA). The relative expression of each gene was determined based on comparative delta-delta CT method ($\Delta\Delta CT$) ⁹³. Statistical significance analysis of expression values of candidate genes was performed using Data Analysis Toolkit in Excel based on one-way ANOVA (P < 0.05) [Supplementary Material 2].

Conclusion

In the present study, as a result of comprehensive network analysis we short listed 12 candidate genes *HK2*, *EXPA1*, *BAM1* (*LRR*), *DELLA* (*RG1*), *GID1C*, *GH3.8*, *ZEP*, *ACLB2*, *LFY*, *MADS16*, *MADS29* and *AGL8* associated to parthenocarpy. By probing the interaction of validated candidate genes we hypothesized a hormone mediated pathway involved in parthenocarpy. Negative regulation of MADS boxTFs (*AGL8* and *MADS16*)together with *DELLA*, *GH3.8* and the proposed pathway of GA mediated auxin induced seedless fruit formation in our previous work has now been supported through their

expression analysis in cvs. C4, CVR and PL. in natural parthenocarp (*Musa* spp.). In view of the fact that interaction of *GID1C* with *DELLA* is a prerequisite for seed formation, drastic reduction *GID1C* together with moderate expression of *DELLA* might be the reason behind seed set in CVR upon pollination. The significance of cytokinin mediated CLV-WUSHEL signaling pathway in parthenocarp is exposed through the interaction of *BAM1*, *IAA (GH3.8)* and *HK2* in PA-PPI network and expression of these genes in *Musa* spp affirmed their role in natural parthenocarp. However, expression of *CLV1* needs validation for further confirmation. Besides, the variation in the expression of *EXPA1*, *ACLB2*, *MADS29*, *LFY* in natural and artificial parthenocarp, emphasize the need for further validation in other diploid, triploid and tetraploid accessions of different genomic combinations. In addition, transcriptome analysis of seeded, poor and no seed setting accessions would enlighten the mechanism of natural parthenocarp. In a nutshell, we suggest that genes *MADS16*, *AGL8*, *DELLA*, *GID1C*, *IAA (GH3.8)*, *HK2*, *BAM1* and *CLV1* could be the possible target genes for manipulation of seeded accessionstoparthenocarp.

Declarations

Author information

Backiyarani Suthanthiram, Sasikala Rajendran, Sharmiladevi Simeon, and Uma Subbaraya

Affiliation

ICAR-National Research Centre for Banana, Thogamalai Road, Thayanur Post, Tiruchirapalli 620 102, Tamil Nadu.

Contributions

US and SB conceived, design the work and provided guidance in manuscript preparation; SB involved in thematic guidance to bring out the work layout and in manuscript preparation; SR constructed PPI network and identified candidate genes through *in-silico* approaches and SS validate candidate genes using qRT-PCR. SR, SB and SS interpreted the results. SR, SB and SS wrote the manuscript paper with input from all authors.

Corresponding author

Correspondence to Uma Subbaraya

Ethics declarations

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

This study was supported by DBT-North east project entitled “Consortium for managing Indian banana genetic Resources – DBT-NER/AGRI/33/2016(Application Number-90)” India.

Acknowledgement

We express our sincere gratitude to the Director, ICAR-National Research Centre for Banana, India for her support in writing this Research article. Our special thanks to Dr.P.Durai,Assistant Chief Tech.Officer ICAR-NRCB, Trichy, Sathish kumar P for their support in banana ovary sample collections. We also thank Ms. Rajeshwari M and Ms. Sindhuja S, project students from Tamil Nadu Agricultural University, Coimbatore for their help in preliminary study of the work.

References

1. Dhatt, A.S. & Kaur, G. Parthenocarp: A potential trait to exploit in vegetable crops: A review. *Agricultural Reviews*. 37,300-308.doi: 10.18805/ag.v37i4.6460.(2016).
2. Gillaspay, G., Ben-David, H., & Gruissem, W. Fruits: A Developmental Perspective. *Plant Cell*. 5:1439-1451.(1993).
3. Sykes, S.R& Lewis, S. Comparing Imperial mandarin and Silverhill Satsuma mandarins as seed parents in a breeding program aimed at developing new seedless Citrus cultivars for Australia. *Australian Journal of Experimental Agriculture* 36,731-38. (1996).
4. Goetz, M., Vivian-Smith, A., Johnson, S.D. & Koltunow, A.M. AUXIN RESPONSE FACTOR8 Is a Negative Regulator of Fruit Initiation in *Arabidopsis*. *Plant Cell*. 18,1873-86.doi: 10.2307/20076744.(2006).
5. Yoshida, T., Matsunaga, S. & Saito, T. Inheritance of parthenocarpic character in eggplant. *J Jpn Soc Hortic Sci*, 67,257.(1998).
6. Kuno, S. & Yabe, K. Genetic analysis of parthenocarp and spineless in the F2 segregating. *Research Bulletin of the Aichi-ken Agricultural Research Center* (Japan) 37, 29-33.(2005).

7. Tiwari, A., Vivian-Smith, A., Roeland, E., Myckel, V.E.J., HabetsXue, L.B., Offringa, R. & Heuvelink, E.P. Parthenocarpic potential in *Capsicum annuum* is enhanced by carpelloid structures and controlled by a single recessive gene. *BMC Plant Biol.* 143,1-14.doi: 10.1186/1471-2229-11-143. (2011).
8. Fos, M., Proano, K., Nuez, F. & Garcia-Martinez, J.L. Role of gibberellins in parthenocarpic fruit development induced by the genetic system pat-3/pat-4 in tomato. *Physiol Plant.*111,545-50. doi:10.1034/j.1399-3054.2001.1110416.x.(2001).
9. Gorguet,, van Heusden, A.W. &Lindhout, P. Parthenocarpic Fruit Development in Tomato. *Plant Biol.* 7,131-139. doi:10.1055/s-2005-837494. (2005).
10. Prohens, J., ruiz J.J&Nuez F. The inheritance of parthenocarp and associated traits in pepino. *Hort. Sci.* 123:376-380. (1998)
11. Kim, I.S., Okubo, H& Fujieda, K. Genetic and hormonal control of parthenocarp in cucumber (*Cucumis sativus* L). *Fat. Agr, Kyushu Univ.* 36,173–81.(1992).
12. Pike, L.M &Peterson, C.E. Inheritance of parthenocarp in the cucumber (*Cucumis sativus*). *Euphytica.* 18,101-5.(1969).
13. Yan, L.Y., Lou, L.N., Feng, Z.H., Lou, Q.F., Chen, X.L. & Li, J.F. *Ying Yong Sheng Tai Xue Bao.* 21,61-66.(2010).
14. Ortiz, R& Vuylsteke, D. Effect of the parthenocarp gene P1 and ploidy on fruit and bunch traits of plantain-banana hybrids. 75:460-465. doi:10.1038/hdy.1995.162. (1995).
15. Serrani, J.C., Ruiz-Rivero, O., Fos, M. &Garcia-Martinez, J.L. Auxin induced fruit set in tomato is mediated in part by gibberellins. *Plant J.* 56:922-34. (2008).
16. Dorcey, E., Urbez, C., Blazquez, M.A., Carbonell, J. & Perez-Amador, M.A. Fertilization-dependent auxin response in ovules triggers fruit development through modulation of gibberellin metabolism in *Arabidopsis*. *Plant J.* 58,318-32.doi:10.1111/j.1365-313X.2008.03781.x.(2009).
17. Pak, H.Y. Effects of plant growth regulators on parthenocarpic fruit development in watermelon (*Citrullus vulgaris*). *J Korean Soc Hort Sci.* 34,167-72.(1993).
18. Hu, J., Israeli, A., Ori, N. & Sun, T.P. The Interaction between DELLA and ARF/IAA Mediates Crosstalk between Gibberellin and Auxin Signaling to Control Fruit Initiation in Tomato.*Plant Cell.* 30, 1710-1728.(2018).
19. Schwabe, W.W. & Mills, J.J. Hormones and parthenocarpic fruit set: A literature survey. *Abstr.*51,661–698. 1981.
20. Tsao,T.plant growth substances:role in fertilization and sex expression. (ed.SKoog, F.)345-348 (spring –verlag,1980).
21. Vivian-Smith, A. & Koltunow A.M. genetic analysis of growth reguator -induced parthenocarp in arabidopsis. 121, 437-452DOI: <https://doi.org/10.1104/pp.121.2.437>. (1999)
22. Serrani, J.C., Sanjuán, R., Ruiz-Rivero, O., Fos, M. & García-Martínez, J.L. Gibberellin Regulation of Fruit Set and Growth in Tomato. *Plant physiology.* 145,246-257. doi.org/10.1104/pp.107.098335.(2007).
23. Chen, X., Zhang, M., Tan, J., Huang, S., Wang, C. & Zhang, H. Comparative transcriptome analysis provides insights into molecular mechanisms for parthenocarpic fruit development in eggplant (*Solanum melongena*). *PLoS ONE.*12:e0179491. doi.org/10.1371/journal.pone.0179491. (2017).
24. Zhang, S., Shi, Q., Albrecht, U., Shatters, Jr R.G., Stange, R., McCollum, G., Zhang, S., Fan, C & Stover, E. Comparative transcriptome analysis during early fruit development between three seedy citrus genotypes and their seedless mutants. *Horti Res* 4, 17041.doi.org/10.1038/hortres.2017.41. (2017).
25. Liu, W., Chen, M., Bai, , Zhuang, Z., Fan, C., Jiang, N., Zhao, J., Ma, S& Xiang. X. Comprehensive transcriptomics and proteomics analyses of pollinated and parthenocarpic litchi (*Litchi chinensis* Sonn.) fruits during early development. *Sci Rep.*7,5401.(2017). doi.org/10.1038/s41598-017-05724-z.
26. Somyong, S., Walayaporn, K., Jomchai, N., Naktang, C., Yodyingyong, T., Phumichai, C., Pootakham, W. & Tangphatsomruang, S. Transcriptome analysis of oil palm inflorescences revealed candidate genes for an auxin signaling pathway involved in parthenocarp. *PeerJ.* 6:e5975. [org/10.7717/peerj.5975](https://doi.org/10.7717/peerj.5975). (2018).
27. Liu, J., Zhai, R., Liu, F., Zhao, Y., Wang, H., Liu, L., Yang, C., Wang, Z., Ma, F& Xu, L. Melatonin Induces Parthenocarp by Regulating Genes in Gibberellin Pathways of 'Starkrimson' Pear (*Pyrus communis*). *Front. Plant. Sci.* 9, 946. doi: 10.3389/fpls.2018.00946.(2018).
28. Galimba, K.D., Bullock, D.G., Dardick, C., Liu, Z., & Callahan, A.M. Gibberellic acid induced parthenocarpic 'Honeycrisp' apples (*Malus domestica*) exhibit reduced ovary width and lower acidity. *Hortic Res.* 6,41.doi:10.1038/s41438-019-0124-8. (2019)

29. Simmonds, N.M. The Development of the Banana Fruit. *J Experimental Botany* 4:87–105. doi.org/10.1093/jxb/4.1.87. (1953).
30. Ortiz, R & Vuylsteke, D. Inheritance of black sigatoka resistance and fruit parthenocarp in the triploid AAB plantain. *Agronomy Abstracts, Madison, WI*, 109, 07. (1992).
31. Simmonds, N. The Evolution of the Bananas. (ed. London) 170. (Longmans, 1962).
32. Dodds, K.S. & Simmond, N.W. Sterility and parthenocarp in diploid hybrids of *musa*. *Heredity*. 2,101-117. doi:10.1038/hdy.1948.6. (1948).
33. Chen, X., Zhang, M., Tan, J., Huang, S., Wang, C. & Zhang, H. Comparative transcriptome analysis provides insights into molecular mechanisms for parthenocarpic fruit development in eggplant (*Solanum melongena*). *PLoS ONE* 12:e0179491. doi.org/10.1371/journal.pone.0179491. (2017).
34. Walhout, A.J., Sordella, R., Lu, X., Hartley, J.L., Temple, G.F., Brasch, M.A., Thierry-Mieg, N. & Vidal, M. Protein interaction mapping in *C. elegans* using proteins involved in vulval deveScience. 287,116–122. (2000).
35. Yadav, S.S & Chouhan, U. Analysis of Protein-Protein Interaction Network of Laminopathy Based on Topological Properties. *BMC Syst Biol*. 11,2. (2018). [org/10.13005/bpj/1470](https://doi.org/10.13005/bpj/1470).
36. Li, S., Choi, K.P & Wu, T. Degree distribution of large networks generated by the partial duplication model network model. *Chaos*. 476, 94-108. doi: 10.1016/J.TCS.2012.12.045. (2013).
37. Martí, C., Orzáez, D., Ellul, P., Moreno, V., Carbonell, J. & Granell, A. Silencing of DELLA induces facultative parthenocarp in tomato fruits. *Plant J*. 52,865-876. doi.org/10.1111/j.1365-313X.2007.03282.x. (2007).
38. Du, L., Bao, C., Hu, T., Zhu, Q., Hu, H., He, Q. & Mao, W. SmARF8, a transcription factor involved in parthenocarp in eggplant. *Mol Genet Genomics*. 291, 93-105. (2015). doi: 10.1007/s00438-015-1088-5.
39. Cong, L., Yue, R., Wang, H., Liu, J., Zhai, R., Yang, J., Wu, M., Si, M., Zhang, H., Yang, C., Xu, L., & Wang, Z. 2,4-D-induced parthenocarp in pear is mediated by enhancement of GA4 biosynthesis. *Physiol Plant*. 166, 812-820. (2018).
40. Liu, S.Y., Zhang, Y.W., Feng, Q.S., Qin, L., Pan, C.T. & Lamin-Samu, A.T. Tomato AUXIN RESPONSE FACTOR 5 regulates fruit set and development via the mediation of auxin and gibberellin signaling. *Sci Rep*. 8,2971. (2018) .doi: 10.1038/s41598-018-21315-y.
41. Backiyarani, S., Uma, S., Sarawathi, M.S., Saravanakumar, A.S. & Chandrasekar, A. Transcriptome analysis of banana (*Musa balbisiana*) based on next-generation sequencing technology. *Turk J Agric For*. 39,705-717. doi:10.3906/tar-1406-171. (2015).
42. Acencio, M. L. & Lemke, N. Towards the prediction of essential genes by integration of network topology, cellular localization and biological process information. *BMC Bioinformatics*. 10,290. (2009).
43. Fraser, H.B., Hirsh, A.E., Steinmetz, L.M., Scharfe, C. & Feldman, W.M. Evolutionary Rate in the Protein Interaction Network. *Science*. 296,750–752. doi:10.1126/science.1068696. (2002).
44. Batada, N.N., Hurst, L.D., & Tyers, M. Evolutionary and Physiological Importance of Hub Proteins. *PLoS Comput. Biol*. 2,0748-0756. (2006).
45. Sharp, P.M. Determinants of DNA sequence divergence between *Escherichia coli* and *Salmonella typhimurium*: Codon usage, map position, and concerted evolution. *J Mol Evol*. 33,23-33. (1991).
46. Rocha, E.P.C. & Danchin, A. An Analysis of Determinants of Amino Acids Substitution Rates in Bacterial Proteins. *Mol Biol Evol*. 21,108–16. doi: 10.1093/molbev/msh004. (2004).
47. Li, G., Li, M., Wang, J., Wu, J., Wu, F.X & Pan, Y. Predicting essential proteins based on subcellular localization, orthology and PPI networks. *BMC Bioinformatics*. 17, 279. doi: 10.1186/s12859-016-1115-5. (2016).
48. Fang, G., Bhardwaj, N., Robilotto, R. & Gerstein, M.B. Getting started in gene Orthology and functional analysis. *PLoS Computer Biology*. 6,6:e1000703. doi: 10.1371/journal.pcbi.1000703. (2010).
49. Mazzucato, A., Taddei, A.R. & Soressi, G.P. The parthenocarpic fruit (pat) mutant of tomato (*Lycopersicon esculentum* Mill.) sets seedless fruits and has aberrant anther and ovule development. *Development*. 125,107-14. (1998).
50. Takisawa, R., Maruyama, T., Nakazaki, T., Kitajima, A. & Kitajima, A. Parthenocarp in the tomato (*Solanum lycopersicum* L.) cultivar 'MPK-1' is controlled by a novel parthenocarpic gene. *Hort J*. 86,487–92. (2017).
51. Pien, S & Grossniklaus, U. Polycomb group and trithorax group proteins in *Arabidopsis*. *Biochim Biophys Acta*. 1769,375-82. (2007).

52. Ohad, N., Yadegari, R., Margossian, L., Hannon, M., Michaeli, D., Harada, J., Goldberg, R.B. & Fischer, R.L. Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell*. 11,407-16. doi: 10.1105/tpc.11.3.407. (1999)..
53. Luo, M., Bilodeau, P., Koltunow, A., Dennis, E.S., Peacock, W.J & Chaudhury, A.M. Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA*. 96, 296–301. (1999).
54. Grossniklaus, U., Vielle-Calzada, J.P., Hoepfner, M.A. & Gagliano, W.B. Maternal Control of Embryogenesis by MEDEA, a Polycomb Group Gene in *Arabidopsis*. *Science* 280,446-450. (1998). doi:10.1126/science.280.5362.446.
55. Joldersma, D. & Liu, Z. The making of virgin fruit: the molecular and genetic basis of parthenocarpy. *J Exp Bot*. 6,955–962. (2018). doi:10.1093/jxb/erx446.
56. Hao, S., Ariizumi, T. & Ezura, H. SEXUAL STERILITY is Essential for Both Male and Female Gametogenesis in Tomato. *Plant Cell Physiol*. 58,22-34. doi.org/10.1093/pcp/pcw214.(2017).
57. Kumar N. Problems and prospects of banana breeding in India. *J Hort Sci*. 1,77-94. (2006).
58. DeYoung, B.J., Bickle, K.L., Schrage, K.J., Muskett, P., Patel, K. & Clark, S.E. The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in *Arabidopsis*. *Plant J*.45,1-16. doi: 10.1111/j.1365-313X.2005.02592.x. (2005).
59. Batte, M., Swennen, R., Uwimana, B., Akech, V., Brown, A., Tumuhimbise, R., Hovmalm, H.P. & Geleta, M, Ortiz R. Crossbreeding East African Highland Bananas: Lessons Learnt Relevant to the Botany of the Crop After 21 Years of Genetic Enhancement. *Front.Plant Sci*. 10,81. <https://doi.org/10.3389/fpls.2019.00081>. (2019).
60. Uma, S. & Arun, K. Understanding the Diversity and Reproductive Biology of Banana – for Improvement through Basic Research. *Acta horticulturae* (abstract). (2016).
61. Uma, S., Sasikala, R., Sharmiladevi, S., Backiyarani, S & Saraswathi, M.S. Unravelling the regulatory network of transcription factors in parthenocarpy. *Sci Horti*. 261,108920. doi.org/10.1016/j.scienta.2019.108920. (2020).
62. Balbi, V. & Lomax, T.L. Regulation of early tomato fruit development by the diageotropica gene. *Plant Physiol* 131,186–197.(2003).
63. Beyer, E.M. & Quebedeaux, B. Parthenocarpy in cucumber: mechanism of action of auxin transport inhibitors. *J. Am. Soc. Hortic. Sci*.99, 385–390.(1974).
64. Peng, J. Gibberellin regulates *Arabidopsis* seed germination via RGL2, a GAI/RGA-like gene whose expression is up-regulated following imbibition. *Genes Dev*. 16, 646-658.doi: 10.1101/gad.969002. (2002).
65. Dill, A., Jung, H.S & Sun, T.P. The DELLA motif is essential for gibberellin-induced degradation of RGA. *Proc Natl Acad Sci USA*. 98,14162–14167. (2001).doi: 10.1073/pnas.251534098.
66. Wen, C.K. & Chang, C. *Arabidopsis* RGL1 encodes a negative regulator of gibberellin responses. *Plant Cell*.114,87-100. (2002).
67. Lee, S., Cheng, H., King, K.E., Wang, W., He, Y., Hussain, A., Lo, J., Harberd, N.P& Peng, J. Gibberellin regulates *Arabidopsis* seed germination via RGL2, a GAI/RGA-like gene whose expression is up-regulated following imbibition. *Genes Develop*. 16: 646–658.(2002).
68. Shinozaki, Y., Ezura, K., Hu, J., Okabe, Y., Benard, C., Prodhomme, D., Gibon, Y., Sun, T.P., Ezura, H & Ariizumi, T. Identification and functional study of a mild allele of SIDEELLA gene conferring the potential for improved yield in tomato. *Sci Rep*. 8,12043.doi: 10.1038/s41598-018-30502-w. (2018).
69. Gallego-Giraldo, C., Hu, J., Urbez, C., Gomez, M.D., Sun, T.P. & Perez-Amador, A.M. Role of the gibberellin receptors GID1 during fruit-set in *Arabidopsis*. *Plant J*. (2014). 79:1020–1032.doi:10.1111/tpj.12603.
70. Qin, Q., Wang, W., Guo, X., Yue, J., Huang, Y., Xu, X., Li, J. & Hou, S. *Arabidopsis* DELLA Protein Degradation Is Controlled by a Type-One Protein Phosphatase, TOPP4. *PLoS Genet*. 10,e1004464. doi:10.1371/journal.pgen.100446. (2014).
71. Ding X, Cao Y, Huang L, Zhao J, Xu C, Li X, et al. Activation of the indole-3-acetic acid-amido synthetase GH3-8 suppresses expansin expression and promote salicylate- and jasmonate-independent basal immunity in rice. *Plant Cell*. 2008; 20: 228–240. pmid:18192436.
72. Fu, J., Yu, H., Li, X., Xiao, J. & Wang, S. Rice GH3 gene family: regulators of growth and development. *Plant Signal Behav*.6,570-574. (2011).doi:10.4161/psb.6.4.14947.
73. Ariizumi, T., Shinozaki, Y. & Ezura, H. Genes that influence yield in tomato. *Breed Sci* 63:3-13. <https://doi.org/10.1270/jsbbs.63.3>. (2013).

74. Tang, N., Deng, W., Hu, G., Hu, N. & Li, Z. Transcriptome profiling reveals the regulatory mechanism underlying pollination dependent and parthenocarpic fruit set mainly mediated by auxin and gibberellin. *PLoS One*. 10,e0125355. doi:10.1371/journal.pone.0125355. (2015).
75. Marowa, P., Ding, A. & Kong, Y. Expansins: roles in plant growth and potential applications in crop improvement. *Plant Cell Rep.* 35,949-965. doi:10.1007/s00299-016-1948-4. (2016).
76. Riefler, M., Novak, O., Strnad, M., and Schmülling, T. *Arabidopsis* cytokinin receptor mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *Plant Cell*. 18, 40-54. doi:10.1105/tpc.105.037796. (2006).
77. Sardos, J., Rouard, M., Hueber, Y., Cenci, A., Hyma, K.E., van den Houwe, I., Hribova, E., Courtois, B. & Roux, N.A. Genome-Wide Association Study on the Seedless Phenotype in Banana (*Musa* spp.) Reveals the Potential of a Selected Panel to Detect Candidate Genes in a Vegetatively Propagated Crop. *PLoS ONE*. 11,e0154448. doi.org/10.1371/journal.pone.0154448. (2016).
78. Pomares-Viciania, T., Río-Celestino, M.D., Román, B., Die, J., Pico, B. & Gómez, P. First RNA-seq approach to study fruit set and parthenocarpy in zucchini (*Cucurbita pepo* L.). *BMC Plant Biol.*19,61. doi.org/10.1186/s12870-019-1632-2. (2019).
79. Szumlanski, A.L. & Nielsen, E. The Rab GTPase RabA4d regulates pollen tube tip growth in *Arabidopsis thaliana*. *Plant Cell* 21:526-544. (2009).
80. Peng, H., Zhao, J. & Neff, M.M. ATAF2 integrates *Arabidopsis* brassinosteroid inactivation and seedling photomorphogenesis. *Dev.* 142,4129-413. (2015).
81. Kang, B.H, Busse, J.S, Dickey, C., Rancour, D.M & Bednarek, S.Y. The *Arabidopsis* cell plate-associated dynamin-like protein, ADL1Ap, is required for multiple stages of plant growth and development. *Plant Physiol.* 126:47-68. doi: 10.2307/4279865. (2001)
82. Pan, L., Lv, S., Yang, N., Lv, Y., Liu, Z., Wu, J & Wang, G. The Multifunction of CLAVATA2 in Plant Development and Immunity. *Front. Plant Sci.*7,1573. (2016). doi.org/10.3389/fpls.2016.01573.
83. Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F.X., Juergens, G. & Laux, T. The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* 100:635-644. (2000).
84. Parcy, F., Nilsson, O., Busch, M.A., Lee, I. & Weigel, D. A genetic framework for floral patterning. *Nature*. 395,561–566. (1998).
85. Busch, M.A., Bomblies, K., & Weigel, D. Activation of a Floral Homeotic Gene in *Arabidopsis*. *Science*.285,585–587. (1999).
86. Teo ZWN, Zhou W, Shen L (2019) Dissecting the function of MADS-box transcription factors in orchid reproductive development. *Front Plant Sci* 10: 1474
87. Martinelli, F., Uratsu, S.L., Reagan, R.L., Chen, Y., Tricoli, D., Fiehn, O., Rocke, D.M., Gasser, C.S. & Dandekar, A.M. Gene regulation in parthenocarpic tomato fruit. *J Exp Bot.*60,3873-90. doi: 10.1093/jxb/erp227. (2009).
88. Droc, G., Lariviae, D., Guignon, V., Yahiaoui, N., This, D., garsmeur, O., Dereeper, A., Hamelin, C., Argout, X., Dufayard, J., Lehgelle, J., Bauens, F., Cenci, A., Pitollat, B., D'Hont, A., Ruiz, M., Rouard, M. & Bocs, S. The Banana Genome Hub. *Databasedoi*:10.1093/database/bat035. (2012).
89. Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., Jensen, L.J. & Mering, C.V. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45,D362-D368. (2017). doi:10.1093/nar/gkw937.
90. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, T.J., Ramage, D., Amin, N., Schwikowski B. & Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13,2498-504 doi: 10.1101/gr.1239303. (2003).
91. Cline, M.S. et al. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* 2:2366–2382. (2007).
92. Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M & Robles, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*.21, 3674-3676..doi.org/10.1093/bioinformatics/bti610. (2005).
93. Livak, J.K & Schmittgen, D.T. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. *Methods* 25,402-408. (2001).doi:10.1006/meth.2001.1262.

Figures

Figure 1

MCODE clusters in PA-PPI network. a)Cluster 1: Comprises of 11 genes belongs to hormone-mediated pathway b)Cluster 2: Comprises of 7 genes belongs to Chromosome segregation c)Cluster 3: Comprises of 9 genes belongs to Carbohydrate metabolic process.

Figure 2

Ovary image of Musa spp. (A) Calcutta 4 (IC no.0642) seeded ovary, (B) cv.Rose (CVR) (IC no. 0638) parthenocarpic accession, upon pollination seed set occurs and (C) Pisang Lilin (PL) (IC no. 0195) parthenocarpic accession, very rare to set seed upon pollination. Arrows indicate location of seeds in seeded accessions (Musa spp.).

Figure 3

Relative expression of genes in ovary of three banana cultivars (C4, CVR and PL) at three conditions namely Un pollinated (UnP), 24 hrs after pollination (P24) and 48 hrs after pollination (P48). Relative expressions of genes with respect to experimental condition (X-axis) were expressed in Log fold change (Y axis). The mean differences between relative gene expressions were analyzed by ANOVA, $p < 0.05$. P values of the two-way ANOVA are shown in [SupplementaryMaterial 2].

Figure 4

KEGG pathway analysis of PA-PPI Network. Majority of the proteins in the PA-PPI network are involved in plant hormone signal transduction (64%) followed by carbohydrate and ubiquitin mediated proteolysis (9%). Details of genes involved in the respective KEGG pathway are given in Supplementary Table 3.

Figure 5

Putative pathway for natural parthenocarpy in *Musa* spp. A) & B) Increase in auxin, GA and Cytokinin induces parthenocarpic fruit formation by down regulating the expression of MADS box transcription factors (TFs), DELLA (RGA1), HK2 and BAM1. MADS box TFs and RGA1 could act as a focal point in auxin mediated GA response in parthenocarpy and in addition auxin responsive genes are mediated by GA by down-regulation of TFs like ARF, IAA9 (Tang et al., 2015) C) HK2, a cytokinin receptor responsible for cytokinin signaling pathway found to interact with PI4KB1, NAC081 AT2S3 DRP1A that are associated to seed development and confirm its role in seed formation. D) Experimental evidences of interacting partners associated with BAM1 suggested the possible role of CLV-WUS signaling pathway in parthenocarpy. Down regulation of BAM1 leads to consequent down regulation of CLV1 followed by WUS and AG that finally leads to seedless fruit formation. E) MEA, FIE, CLF and SWN are group of polycomb (PcG) proteins (histone modifying enzymes), predicted to involved in lysine degradation pathway and the end product is the formation of glycine and carnitine. In other way the same proteins are speculated to be involved in epigenetic mechanism of seedless fruit formation. Highlighted genes (in red) in the pathway denotes candidate genes of parthenocarpy; Green color box indicates hormonal and other signaling pathway that are predicted to mediate seedless fruit formation; Red dotted arrow lines are putative protein- protein interactions in PA-PPI network; ↓ symbol indicates down regulated expression of the genes.

Figure 6

General workflow of the study. In 1st stage, genes associated with parthenocarpy were mined through literature search. In 2nd stage, corresponding orthologous genes in banana were retrieved using BLAST. In 3rd stage, PPI network was constructed using STRING and Cytoscape software. In 4th stage, topological, cluster analysis identified candidate genes and proposed putative pathway. In 5th stage, candidate genes were validated in *Musa* spp. through qRT-PCR.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial1.xlsx](#)
- [SupplementaryMaterial2.xlsx](#)
- [SupplementaryFigures.docx](#)
- [SupplementaryTables.docx](#)