Investigating sex expression and modification in dioecious *Coccinia grandis* through an integrated transcriptomic and proteomic approach

A Thesis

Submitted in partial fulfilment of the requirements
for the degree of
Doctor of Philosophy

By

Ravi S. Devani

20123181

CERTIFICATE

Certified that the work incorporated in the thesis entitled **"Investigating sex expression and modification in dioecious *Coccinia grandis* through an integrated transcriptomic and proteomic approach"** submitted by **Mr. Ravi S. Devani** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Dr. Anjan K. Banerjee,

Supervisor

Date: 20th April, 2018

डॉ. अंजन के. बॅनर्जी/Dr. Anjan K. Banerjee
सहयोगी प्राध्यापक एवं उप अध्यक्ष, जीवशास्त्र कार्यक्रम
Associate Prof. & Co-Chair, Biology Program
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान
Indian Institute of Science Education & Research
पुणे / Pune - 411 008, India

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Ravi S. Devani

Reg. No. 20123181

Date: 20/4/2018

ACKNOWLEDGEMENTS

Ravi

20th April 2018

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Ag-H flower buds: Morphologically hermaphrodite flowers from plant treated with silver nitrate

ALSV: Apple Latent Spherical Virus

AVG: Aminoethoxyvinylglycine

BUSCO: Benchmarking Universal Single-Copy Orthologs

CRB-BLAST: Conditional Reciprocal Best BLAST

edgeR: Empirical Analysis of Digital Gene Expression Data in R

eggNOG: evolutionary genealogy of genes: Non-supervised Orthologous Groups

FPKM: Fragments per kb per million reads

GO: Gene ontology

GyM: Gynomonoecious

GyM-H flower buds: Morphologically hermaphrodite flowers from a gynomonoecious plant

HMMER: Tool used for searching sequence databases for sequence homologs

iCODEHOP: COnsensus-DEgenerate Hybrid Oligonucleotide Primers

IDA: Information Dependent Acquisition

KEGG: Kyoto encyclopedia of genes and genomes

LC-MS: Liquid Chromatography–Mass Spectrometry

NGS: Next-generation sequencing

Nr: Non-redundant protein database

PCA-DA: Principal Component Analysis and Discriminant Analysis

PFAM: database of protein families, each represented by multiple sequence alignments and hidden Markov models (HMMs)

PTGS: Post-Transcriptional Gene Silencing

qRT-PCR: Quantitative real-time polymerase chain reaction

RdRp: RNA-dependent RNA polymerase

RIN: RNA integrity number

RISC: RNA-Induced Silencing Complex

RSEM: RNA-Seq by Expectation-Maximization

RT-PCR: Reverse transcription polymerase chain reaction

SWATH-MS: Sequential Window Acquisition of All Theoretical Mass Spectra

Swiss-Prot: Annotated protein sequence database

TrEMBL: Computer-annotated supplement to Swiss-Prot database

TRV: Tobacco Rattle Virus

VIGS: Virus Induced Gene Silencing

SYNOPSIS

Investigating sex expression and modification in dioecious *Coccinia grandis* through an integrated transcriptomic and proteomic approach

Name: Devani Ravi Sureshbhai

Roll number: 20123181

Name of supervisor: Dr. Anjan K. Banerjee

Department: Biology, IISER Pune

Date of registration: 1st August, 2012

Chapter 1: Introduction

Sexual systems in plants are almost overwhelmingly diverse. Strikingly, in contrast to animals, most of the species of flowering plants are co-sexual including hermaphrodite species (bearing perfect bisexual flowers) and monoecious species (bearing unisexual male and female flowers on the same individual). However, around 5-6 % of angiosperm species are dioecious (bearing unisexual male and female flowers on separate individuals) (Yampolsky, 1922, Ming *et al.*, 2011, Renner, 2014). There are 15600 dioecious angiosperms (5–6 % of the total species) in 987 genera (7 % of genera) and 175 families (43 % of families), with somewhere between 871 to 5000 independent origins of dioecy (Renner, 2014). The patchy distribution of these dioecious species over the phylogenetic tree indicates that dioecy might have evolved multiple times in different families independently (Ainsworth, 2000, Ming *et al.*, 2011, Renner, 2014). The evolution of dioecy would require occurrence and establishment of at least two independent mutations that include a male-sterility (M → m) mutation and a female-sterility mutation ($Su^f \rightarrow Su^{Female}$) (Lewis, 1942, Charlesworth and Charlesworth, 1978, Ross, 1978, Charlesworth, 2002, Charlesworth, 2013). Also, these two mutations should be closely linked to prevent the appearance of neuters and hermaphrodites due to their recombination. Eventually, the chromosomes harbouring such mutations will evolve into sex chromosomes. Molecular and genetic mechanisms of sex determination have been studied in many monoecious and dioecious species such as *Silene latifolia* (Hardenack *et al.*, 1994, Lebel-Hardenack *et al.*, 2002, Muyle *et al.*, 2012), *Rumex acetosa* (Kihara and Ono, 1925, Ainsworth *et al.*, 1995, Navajas-Pérez *et al.*,

2005), *Carica papaya* (Liu *et al.*, 2004, Yu *et al.*, 2008, Urasaki *et al.*, 2012), *Asparagus officinalis* (Bracale et al., 1990, Harkess et al., 2015, Tsugama et al., 2017), *Diospyros kaki* (Akagi *et al.*, 2014), *Cucumis sativus* (Boualem *et al.*, 2009, Guo *et al.*, 2010), *Cucumis melo* (Boualem *et al.*, 2008, Martin *et al.*, 2009, Boualem *et al.*, 2015), *Zea mays* (Chuck *et al.*, 2007), *Spinacia oleracea* (Pfent *et al.*, 2005, Sather *et al.*, 2010), *Thalictrum* (Di Stilio *et al.*, 2005) etc. Since it is believed that evolution of dioecy has happened multiple times in angiosperms independent of each other, the molecular mechanisms of achieving unisexual flower development will most probably vary between different groups of dioecious species (Dellaporta and Calderon-Urrea, 1993).

Coccinia grandis is one such dioecious species having heteromorphic sex chromosomes (Male: 22A+XY & Female: 22A+XX) (Kumar and Viseveshwaraiah, 1952, Bhowmick *et al.*, 2012, Sousa *et al.*, 2013). Y-chromosome of male sex is at least twice the size of the biggest autosome and plays a decisive role in determining maleness (Sousa *et al.*, 2013). *C. grandis* belongs to Cucurbitaceae, a family known for its diverse sexual systems (Kouonon *et al.*, 2009). *Coccinia grandis* has received less attention as a system to understand dioecy compared to *S. latifolia*, *R. acetosa*, *Populus*, etc. Availability of genome sequences for four cucurbit species gives *C. grandis*, an advantage as a system to study dioecy. *C. grandis*, commonly known as ivy gourd, is used as a vegetable and has rich medicinal value.

Since we were interested in establishing *Coccinia grandis* as a model system to understand dioecy, at the beginning, we have carried out detailed morphological, histological and cytological studies with three difference sexual forms of *C. grandis* (Ghadge *et al.*, 2014). In addition to male and female plants, a rare gynomonoecious (GyM) form of *C. grandis* bearing pistillate (GyM-F) and morphologically hermaphrodite flowers (GyM-H) was found during our study (Ghadge *et al.*, 2014). Morphometric analysis showed the presence of staminodes in flowers of female plant and histological study revealed the absence of carpel initials in the flowers of male plant. Though GyM plant had XX sex chromosomes, the development of stamens occurred in hermaphrodite flowers (GyM-H) in addition to carpels. Interestingly, the sex of *C. grandis* female plants was modified upon treatment with $AgNO_3$ which lead to the development of morphologically hermaphrodite flowers (Ag-H) (Ghadge *et al.*, 2014). However, the characteristic development of stamens in hermaphrodite flowers of GyM plant having XX

sex chromosomes indicates that Y-chromosome dependency is somehow bypassed. Interestingly, the pollens of Ag-H and GyM-H flowers were found to be sterile, indicating the probable role of Y-chromosome in governing pollen fertility (Ghadge *et al.*, 2014). Our experimental findings together with all other previous chromosomal and molecular cytogenetical data strongly support the view that *C. grandis* could be used as a potential model system to study sex expression in dioecious flowering plant.

To understand sex expression and modification in dioecious *C. grandis*, following objectives were selected for the present investigation:

1. Candidate gene approach and comparative transcriptomics to identify novel players involved in sex expression and modification of *C. grandis*.
2. Proteomic characterization of *C. grandis* flower buds in order to understand unisexual flower development and $AgNO_3$ mediated sex modification.
3. Standardization of virus induced gene silencing (VIGS) in *C. grandis* for functional characterization of key differentially expressed genes.

Chapter 2: Candidate gene approach and comparative transcriptomics to identify novel players involved in sex expression and modification of *C. grandis*.

Due to the unavailability of genome sequence and the slow pace at which sex-linked genes are identified, sex expression and modification in *C. grandis* are not well understood. In the beginning, a candidate gene approach was taken to isolate *C. grandis* homologs for sex-determining genes of other species using degenerate primers. Homologs of candidate genes *PI*, *AG* and *ACS* were isolated from *C. grandis* by degenerate primer approach. Also, *miR172* was isolated from *C. grandis* by stem-loop PCR. Expression analysis of these candidate genes revealed male-biased expression of *CgPI*. Interestingly, *CgACS* expression was approx. 10-fold higher in female buds compared to male. This pattern was similar to the expression profiles of *CmACS-7* (melon) and *CsACS2* (cucumber) indicating a possibility of conserved sex determination mechanisms. Finally, we have carried out a comprehensive RNA-Seq study from early-staged male, female, GyM-H, and Ag-H as well as middle-staged male and GyM-H flower buds (Devani *et al.*, 2017). A *de novo* transcriptome was assembled using Trinity and annotated by BLAST2GO and Trinotate pipelines. The assembled transcriptome consisted of 467,233 'Trinity Transcripts' clustering into 78,860 'Trinity Genes'. Female_Early_vs_Male_Early,

Ag_Early_vs_Female_Early, and GyM-H_Middle_vs_Male_Middle comparisons exhibited 35,694, 3574, and 14,954 differentially expressed transcripts respectively. Further, qRT-PCR analysis of selected candidate genes validated digital gene expression profiling results. Interestingly, ethylene response-related genes were found to be upregulated in female buds compared to male buds (Devani *et al.*, 2017). Also, we observed that AgNO$_3$ treatment suppressed ethylene responses in Ag-H flowers by downregulation of ethylene-responsive transcription factors leading to stamen development (Devani *et al.*, 2017). Further, GO terms related to stamen development were enriched in early-staged male, GyM-H, and Ag-H buds compared to female buds supporting the fact that stamen growth gets arrested in female flowers. Suppression of ethylene responses in both male and Ag-H compared to female buds suggests a probable role of ethylene in stamen suppression similar to monoecious cucurbits such as melon and cucumber. Also, pollen fertility associated GO terms were depleted in middle-staged GyM-H buds compared to male buds indicating the necessity of Y-chromosome for pollen fertility (Devani *et al.*, 2017). Overall, this study would enable identification of new sex-biased genes for further investigation of stamen arrest, pollen fertility, and AgNO$_3$-mediated sex modification.

Chapter 3: Proteomic characterization of *C. grandis* flower buds in order to understand unisexual flower development and AgNO$_3$ mediated sex modification.

To identify sex-related differences, total proteome from male, female, GyM-H and Ag-H flower buds at early and middle stages of development were analysed by label-free proteomics (Devani *et al.*, 2018: Under review). 3385 proteins were detected (FDR ≤ 1) using *C. grandis* flower bud transcriptome database. SWATH-MS based comparative abundance analysis of 2262 proteins between Female_Early vs Male_Early, Ag_Early vs Female_Early, GyM-H_Middle vs Male_Middle and Ag_Middle vs Male_Middle identified 644, 849, 669 and 591 differentially expressed proteins, respectively. Upregulation of ethylene biosynthesis-related proteins in female buds compared to male buds indicated a conserved stamen arrest mechanism similar to monoecious cucurbits. AgNO$_3$ treatment induced proteins related to pollen development in Ag-H buds. However, few proteins governing pollen germination and tube growth were highly expressed in male buds compared to Ag-H and GyM-H buds indicating the role of Y-chromosome in governing pollen fertility. Overall, our findings will help in identification of key molecules governing dioecy and unisexual flower development in Cucurbitaceae.

Chapter 4: Standardization of virus induced gene silencing (VIGS) in *C. grandis* for functional characterization of key differentially expressed genes.

Our transcriptomic and proteomic studies have identified many interesting sex-biased genes/proteins which might govern the stamen inhibition in female flowers and pollen fertility in male flowers of *C. grandis*. However, lack of an efficient *in vitro* regeneration and genetic transformation tool for *C. grandis* was a major limitation to study the function of the sex-biased genes identified in our study. Initially, we attempted to achieve *in vitro* regeneration and transformation of *C. grandis* in our lab. *Agrobacterium*-mediated infection was achieved in the leaf and tendril explants of *C. grandis*, which was confirmed by GUS staining. However, we failed to regenerate shoots from these transformed calli even after testing various growth hormone concentrations. Hence, we chose to standardize virus induced gene silencing tool as an alternative strategy for gene knockdown in *C. grandis* using tobacco rattle virus (TRV) and apple latent spherical virus (ALSV). We have standardized ALSV-VIGS system for *C. grandis* in order to understand gene function (Devani *et al.*, 2018: Under preparation). Effective silencing of *CgPDS* gene in *C. grandis* plants using ALSV-VIGS system was achieved (Devani *et al.*, 2018: Under preparation). Attempts are also ongoing to study candidate DE genes from our transcriptomics and proteomics approach using ALSV-VIGS in order to understand their role in sex expression and modification.

Summary

Out literature review suggested that *C. grandis* has a unique advantage as a model system to study dioecy as it belongs to Cucurbitaceae family, where the sex determination mechanisms for many of its monoecious relatives have been well understood. Also, availability of genome sequences for multiple cucurbits would help in identification and annotation of genes in *C. grandis*. In this study, we began with detailed morphological, histological and cytological studies with three different sexual forms of *C. grandis* (Ghadge *et al.*, 2014). We also identified a rare gynomonoecious form of *C. grandis* (22+XX) that bears pistillate (GyM-F) and morphologically hermaphrodite flowers (GyM-H). Interestingly, we observed that application of silver nitrate (AgNO₃) on female plants induces stamen development leading to the formation of morphologically hermaphrodite flowers (Ag-H) despite the absence of Y-chromosome. An omics approach was taken to identify interesting candidate genes that might govern the sex

expression. RNA from early-staged male, female, GyM-H and Ag-H as well as middle-staged male and GyM-H flower buds was sequenced on Illumina platform. A *de novo* transcriptome was assembled and annotated. Female_Early vs Male_Early, Ag_Early vs Female_Early, and GyM-H_Middle vs Male_Middle comparisons exhibited 35694, 3574, and 14954 differentially expressed transcripts respectively (Devani *et al.*, 2017). Simultaneously, total proteome from male, female, GyM-H and Ag-H flower buds at early and middle stages of development were analysed by label-free proteomics (Devani *et al.*, 2018: Under review). 3385 proteins were detected (FDR ≤ 1) using *C. grandis* flower bud transcriptome database. SWATH-MS based comparative abundance analysis of 2262 proteins between Female_Early vs Male_Early, Ag_Early vs Female_Early, GyM-H_Middle vs Male_Middle and Ag_Middle vs Male_Middle identified 644, 849, 669 and 591 differentially expressed proteins, respectively. Ethylene responses were found to be suppressed in both male and Ag-H compared to female buds suggesting a probable role of ethylene in stamen suppression similar to monoecious cucurbits such as melon and cucumber. Also, pollen fertility associated GO-terms were depleted in middle-staged GyM-H buds compared to male buds indicating the necessity of Y-chromosome for pollen fertility. In absence of genetic transformation technique for *C. grandis*, we have standardized ALSV-VIGS system to understand the function of sex-biased genes identified in our transcriptomics and proteomics study. Effective silencing of *CgPDS* gene in *C. grandis* plants was achieved by ALSV-VIGS (Devani *et al.*, 2018: Under preparation). Attempts are ongoing to study candidate DE genes from our transcriptomics and proteomics approach using ALSV-VIGS in order to understand their role in sex expression and modification. Overall, our study has enabled identification of new sex-biased genes and provided a basis for functional characterization to understand their role in stamen arrest, pollen fertility, and AgNO$_3$-mediated sex modification.

List of Publications

- Ghadge A., Karmakar K., **Devani R.S.**, Banerjee J., Mohanasundaram B., Sinha R.K., Sinha S., Banerjee A.K. (2014). Flower development, pollen fertility and sex expression analyses of three sexual phenotypes of *Coccinia grandis*. **BMC Plant Biology.** 14: 325. DOI: 10.1186/s12870-014-0325-0

- **Devani R.S.**, Sinha S., Banerjee J., Sinha R.K., Bendahmane A. and Banerjee A.K. (2017). *De novo* transcriptome assembly from flower buds of dioecious, gynomonoecious

and chemically masculinized female *Coccinia grandis* reveals genes associated with sex expression and modification. **BMC Plant Biology.** DOI:10.1186/s12870-017-1187-z

- **Devani R.S.**, Chrimade T., Sinha S., Bendahmane A., Banerjee A.K.[*] and Banerjee J[*]. (2018). Flower bud proteome reveals modulation of sex-biased proteins potentially associated with sex expression and modification in dioecious *Coccinia grandis*. **Under review.**

- Sinha S., Karmakar K., **Devani R.S.**, Banerjee J., Sinha R.K., Banerjee A.K. (2016). Preparation of Mitotic and Meiotic Metaphase Chromosomes from Young Leaves and Flower Buds of *Coccinia grandis*. **bio-protocol**. 6(7), e1771. DOI: https://doi.org/10.21769/BioProtoc.1771

- **Devani R.S. *et al.*** (2018). A virus induced gene silencing tool for the study of gene function in dioecious *Coccinia grandis*. **Under preparation.**

References

Ainsworth, C. (2000) Boys and Girls Come Out to Play: The Molecular Biology of Dioecious Plants. *Annals of Botany*, **86**, 211-221.

Ainsworth, C., Crossley, S., Buchanan-Wollaston, V., Thangavelu, M. and Parker, J. (1995) Male and female flowers of the dioecious plant sorrel show different patterns of MADS box gene expression. *The Plant Cell*, **7**, 1583.

Akagi, T., Henry, I.M., Tao, R. and Comai, L. (2014) A Y-chromosome–encoded small RNA acts as a sex determinant in persimmons. *Science*, **346**, 646.

Bhowmick, B.K., Jha, T.B. and Jha, S. (2012) Chromosome analysis in the dioecious cucurbit *Coccinia grandis* (L.) Voigt. *Chromosome Science*, **15**, 9-15.

Boualem, A., Fergany, M., Fernandez, R., Troadec, C., Martin, A., Morin, H., Sari, M.-A., Collin, F., Flowers, J.M., Pitrat, M., Purugganan, M.D., Dogimont, C. and Bendahmane, A. (2008) A Conserved Mutation in an Ethylene Biosynthesis Enzyme Leads to Andromonoecy in Melons. *Science*, **321**, 836.

Boualem, A., Troadec, C., Camps, C., Lemhemdi, A., Morin, H., Sari, M.-A., Fraenkel-Zagouri, R., Kovalski, I., Dogimont, C., Perl-Treves, R. and Bendahmane, A. (2015) A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. *Science*, **350**, 688.

Boualem, A., Troadec, C., Kovalski, I., Sari, M.-A., Perl-Treves, R. and Bendahmane, A. (2009) A Conserved Ethylene Biosynthesis Enzyme Leads to Andromonoecy in Two *Cucumis* Species. *PLOS ONE*, **4**, e6144.

Bracale, M., Galli, M., Falavigna, A. and Soave, C. (1990) Sexual differentiation in *Asparagus officinalis* L. *Sexual Plant Reproduction*, **3**, 23-30.

Charlesworth, B. and Charlesworth, D. (1978) A Model for the Evolution of Dioecy and Gynodioecy. *The American Naturalist*, **112**, 975-997.

Charlesworth, D. (2002) Plant sex determination and sex chromosomes. *Heredity*, **88**, 94.

Charlesworth, D. (2013) Plant sex chromosome evolution. *Journal of Experimental Botany*, **64**, 405-420.

Chuck, G., Meeley, R., Irish, E., Sakai, H. and Hake, S. (2007) The maize *tasselseed4* microRNA controls sex determination and meristem cell fate by targeting *Tasselseed6/indeterminate spikelet1*. *Nature Genetics*, **39**, 1517.

Dellaporta, S.L. and Calderon-Urrea, A. (1993) Sex determination in flowering plants. *The Plant Cell*, **5**, 1241.

Devani, R.S., Sinha, S., Banerjee, J., Sinha, R.K., Bendahmane, A. and Banerjee, A.K. (2017) *De novo* transcriptome assembly from flower buds of dioecious, gynomonoecious and chemically masculinized female *Coccinia grandis* reveals genes associated with sex expression and modification. *BMC Plant Biology*, **17**, 241.

Di Stilio, V.S., Kramer, E.M. and Baum, D.A. (2005) Floral MADS box genes and homeotic gender dimorphism in *Thalictrum dioicum* (Ranunculaceae) – a new model for the study of dioecy. *The Plant Journal*, **41**, 755-766.

Ghadge, A.G., Karmakar, K., Devani, R.S., Banerjee, J., Mohanasundaram, B., Sinha, R.K., Sinha, S. and Banerjee, A.K. (2014) Flower development, pollen fertility and sex expression analyses of three sexual phenotypes of *Coccinia grandis*. *BMC Plant Biology*, **14**, 325.

Guo, S., Zheng, Y., Joung, J.-G., Liu, S., Zhang, Z., Crasta, O.R., Sobral, B.W., Xu, Y., Huang, S. and Fei, Z. (2010) Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. *BMC Genomics*, **11**, 384.

Hardenack, S., Ye, D., Saedler, H. and Grant, S. (1994) Comparison of *MADS box* gene expression in developing male and female flowers of the dioecious plant white campion. *The Plant Cell*, **6**, 1775.

Harkess, A., Mercati, F., Shan, H.-Y., Sunseri, F., Falavigna, A. and Leebens-Mack, J. (2015) Sex-biased gene expression in dioecious garden asparagus (*Asparagus officinalis*). *New Phytologist*, **207**, 883-892.

Kihara, H. and Ono, T. (1925) The sex-chromosomes of *Rumex acetosa*. *Zeitschrift für induktive Abstammungs-und Vererbungslehre*, **39**, 1-7.

Kouonon, L.C., Jacquemart, A.-L., Zoro Bi, A.I., Bertin, P., Baudoin, J.-P. and Dje, Y. (2009) Reproductive biology of the andromonoecious *Cucumis melo* subsp. agrestis (Cucurbitaceae). *Annals of Botany*, **104**, 1129-1139.

Kumar, L. and Viseveshwaraiah, S. (1952) Sex mechanism in *Coccinia indica* Wight and Arn. *Nature*, **170**, 330.

Lebel-Hardenack, S., Hauser, E., Law, T.F., Schmid, J. and Grant, S.R. (2002) Mapping of Sex Determination Loci on the White Campion (*Silene latifolia*) Y Chromosome Using Amplified Fragment Length Polymorphism. *Genetics*, **160**, 717.

Lewis, D. (1942) The Evolution Of Sex In Flowering Plants. *Biological Reviews*, **17**, 46-67.

Liu, Z., Moore, P.H., Ma, H., Ackerman, C.M., Ragiba, M., Yu, Q., Pearl, H.M., Kim, M.S., Charlton, J.W., Stiles, J.I., Zee, F.T., Paterson, A.H. and Ming, R. (2004) A primitive Y-chromosome in papaya marks incipient sex chromosome evolution. *Nature*, **427**, 348.

Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., Pitrat, M., Dogimont, C. and Bendahmane, A. (2009) A transposon-induced epigenetic change leads to sex determination in melon. *Nature*, **461**, 1135.

Ming, R., Bendahmane, A. and Renner, S.S. (2011) Sex Chromosomes in Land Plants. *Annual Review of Plant Biology*, **62**, 485-514.

Muyle, A., Zemp, N., Deschamps, C., Mousset, S., Widmer, A. and Marais, G.A. (2012) Rapid *de novo* evolution of X-chromosome dosage compensation in *Silene latifolia*, a plant with young sex chromosomes. *PLoS biology*, **10**, e1001308.

Navajas-Pérez, R., de la Herrán, R., López González, G., Jamilena, M., Lozano, R., Ruiz Rejón, C., Ruiz Rejón, M. and Garrido-Ramos, M.A. (2005) The Evolution of Reproductive Systems and Sex-Determining Mechanisms Within Rumex (Polygonaceae) Inferred from Nuclear and Chloroplastidial Sequence Data. *Molecular Biology and Evolution*, **22**, 1929-1939.

Pfent, C., Pobursky, K.J., Sather, D.N. and Golenberg, E.M. (2005) Characterization of *SpAPETALA3* and *SpPISTILLATA*, B class floral identity genes in *Spinacia oleracea*, and their relationship to sexual dimorphism. *Development Genes and Evolution*, **215**, 132-142.

Renner, S.S. (2014) The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, **101**, 1588-1596.

Ross, M.D. (1978) The Evolution Of Gynodioecy And Subdioecy. *Evolution*, **32**, 174-188.

Sather, D.N., Jovanovic, M. and Golenberg, E.M. (2010) Functional analysis of B and C class floral organ genes in spinach demonstrates their role in sexual dimorphism. *BMC Plant Biology*, **10**, 46.

Sousa, A., Fuchs, J. and Renner, S.S. (2013) Molecular Cytogenetics (FISH, GISH) of *Coccinia grandis*: A ca. 3 myr-Old Species of Cucurbitaceae with the Largest Y/Autosome Divergence in Flowering Plants. *Cytogenetic and Genome Research*, **139**, 107-118.

Tsugama, D., Matsuyama, K., Ide, M., Hayashi, M., Fujino, K. and Masuda, K. (2017) A putative MYB35 ortholog is a candidate for the sex-determining genes in *Asparagus officinalis*. *Scientific Reports*, **7**, 41497.

Urasaki, N., Tarora, K., Shudo, A., Ueno, H., Tamaki, M., Miyagi, N., Adaniya, S. and Matsumura, H. (2012) Digital Transcriptome Analysis of Putative Sex-Determination Genes in Papaya (*Carica papaya*). *PLOS ONE*, **7**, e40904.

Yampolsky, C. (1922) Distribution of sex forms in the phanerogamic flora. *Bibliotheca Genetica*, **3**, 1-62.

Yu, Q., Navajas-Pérez, R., Tong, E., Robertson, J., Moore, P.H., Paterson, A.H. and Ming, R. (2008) Recent Origin of Dioecious and Gynodioecious Y Chromosomes in Papaya. *Tropical Plant Biology*, **1**, 49-57.

Chapter 1:   Introduction

1.1 Diversity of sexual systems in plants

Sexual systems in plants are almost overwhelmingly diverse. Strikingly, in contrast to animals, most of the species of flowering plants are co-sexual including hermaphrodite species (bearing perfect bisexual flowers) (e.g., *Arabidopsis thaliana*, *Antirrhinum majus*) and monoecious species (bearing unisexual male and female flowers on the same individual) (e.g., *Cucumis sativus*, *Zea mays*, *Cucumis melo*) (Figure 1-1). However, around 5-6 % of angiosperm species are dioecious (bearing unisexual male and female flowers on separate individuals) (e.g., *Silene latifolia*, *Rumex acetosa*) (Yampolsky, 1922, Ming *et al.*, 2011, Renner, 2014) (Figure 1-1). Besides these main sexual states, many other rare intermediate sex forms also exist. For example andromonoecy (male flowers and bisexual flowers on the same individual), gynomonoecy (female flowers and bisexual flowers on the same individual), trimonoecy (male flowers, female flowers, and bisexual flowers on the same individual), androdioecy (male flowers and bisexual flowers on separate individuals), gynodioecy (female flowers and bisexual flowers on separate individuals) and trioecy (or subdioecy) (male flowers, female flowers and bisexual flowers on separate individuals) (Ainsworth *et al.*, 1997b). Dioecism provides a unique opportunity for studying genetic basis of sex determination, where plants bearing male flowers might be genetically distinct from plants bearing female flowers. This kind of mechanism cannot be studied in monoecious system, since each individual though being genetically uniform, produces both type of flowers, *viz*. male and female.
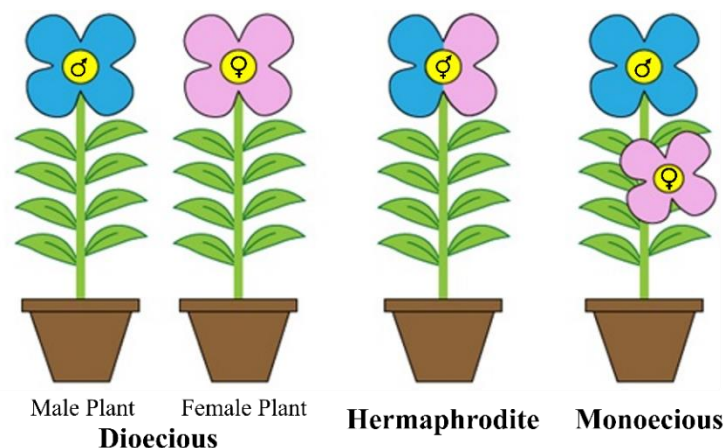
Figure 1-1. **Major sexual states in plants.** Dioecious species have separate individuals bearing male and female flowers. Hermaphrodite species have individuals bearing bisexual flowers. Monoecious species have individuals bearing unisexual male and female flowers on the same individual.

There are 15,600 dioecious angiosperms (5–6% of the total species) in 987 genera (7% of genera) and 175 families (43% of families), with somewhere between 871 to 5000 independent origins of dioecy (Renner, 2014) (Figure 1-2). The patchy distribution of these dioecious species over the phylogenetic tree indicates that dioecy might have evolved multiple times in different families independently (Ainsworth, 2000, Ming *et al.*, 2011). It is observed that dioecy is more common among dicots compared to monocots (Renner and Ricklefs, 1995).
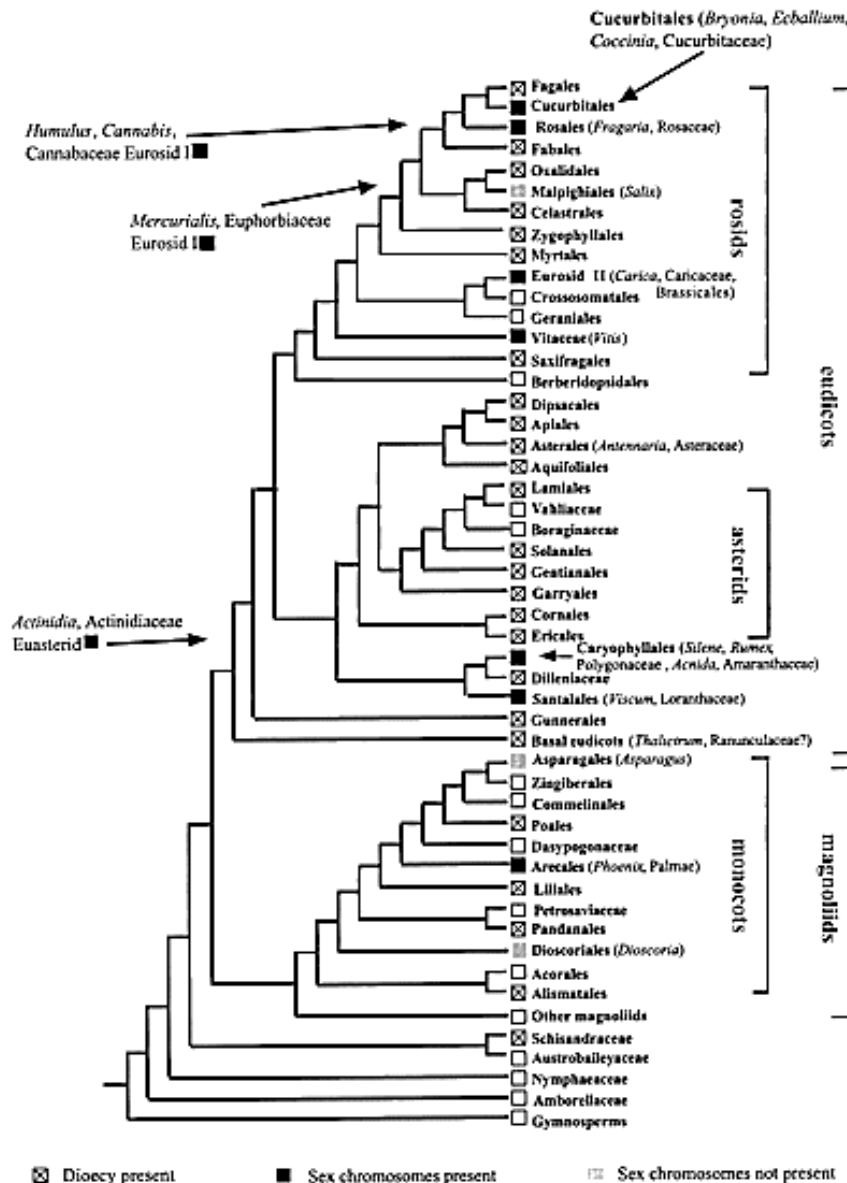
Figure 1-2. **Taxonomic distribution of dioecy and sex chromosomes in angiosperms.** Boxes with X indicate the presence of dioecy. Black boxes indicate the presence of species with heteromorphic sex chromosomes. Grey boxes indicate that sex chromosomes are believed to be absent in the family. Reproduced from Charlesworth (2002).

1.2 Driving forces and evolutionary routes to dioecy

Considering evolution of dioecy, the obvious question that comes to mind is, why would plants evolve separate sexes giving up the reproductive assurance provided by hermaphroditism? Outcrossing advantage logically seems to be the most important factor that would have driven the evolution of dioecious species from hermaphrodite species that suffered from the deleterious effects of inbreeding (Baker, 1959, Lloyd, 1975, Lloyd, 1976, Charlesworth and Charlesworth, 1978). High rates of self-pollination can result in frequent appearance of recessive traits because of more number of homozygous deleterious genes. This results in inbreeding depression i.e. reduced fitness of offspring. Mutant individuals bearing flowers of only one sex would promote outbreeding and hence will be selected over self-pollinating hermaphroditic individuals. In this case, we assumed that the ancestors of dioecious plants are hermaphrodite plants suffering from inbreeding depression. However, this might not be the case with all the hermaphrodite species. Many of the hermaphrodite species have evolved mechanisms other than dioecy to avoid self-pollination and inbreeding depression. These mechanisms include dichogamy (maturation of anthers and stigmas at different times on an otherwise bisexual flower) (Bertin and Christian, 1993) as well as self-incompatibility mechanisms (Nasrallah and Nasrallah, 1993, Newbigin *et al.*, 1993). The evolution of dioecy from such species can be driven by sexual specialization and improved resource allocation (Darwin, 1876, Freeman *et al.*, 1997). Availability of resources for developing pollens and ovule might be limiting in many cases. In such cases, formation of pollens can limit the production of ovules and vice-versa. Sexual specialization (production of either pollen or ovule but not both) can result in better allocation of the available resources. Hence, such sexually specialized mutants bearing either only male or only female flowers get selected over the hermaphrodites. Apart from these forces, many other factors like sexual selection, pollination; and flower and seed predation could also play a role and can even override these forces (Bawa, 1980).

And now, when we know about the forces driving the evolution of dioecy, another question arises is; what are the steps involved in evolution from hermaphroditism towards dioecy? The evolution of dioecy would require occurrence and establishment of at least two independent mutations that include a male-sterility mutation ($M \rightarrow m$) and a female-sterility mutation ($Su^f \rightarrow Su^{Female}$) (Lewis, 1942, Charlesworth and Charlesworth, 1978, Ross, 1978,

Charlesworth, 2002, Charlesworth, 2013) (Figure 1-3). Also, these two mutations should be closely linked to prevent the appearance of neuters and hermaphrodites due to their recombination (Figure 1-3). Simultaneous appearance of both these mutations and hence, the direct evolution of dioecism from hermaphroditism is quite unlikely. The most common evolutionary route is via gynodioecy (Charlesworth and Guttman, 1999). Individuals with male-sterility mutation (M → m) (female plants) will spread in the inbreeding hermaphroditic population if they have equal or higher seed production, resulting in gynodioecious population (Figure 1-3). Then other mutations in hermaphrodite individuals that can improve male functions at the cost of female functions will be favoured because such subhermaphrodites could contribute more genes via pollens than via ovules. Eventually, female functions could get completely suppressed resulting in male individuals. From such a population, dioecy would eventually arise when the two mutations are closely linked.
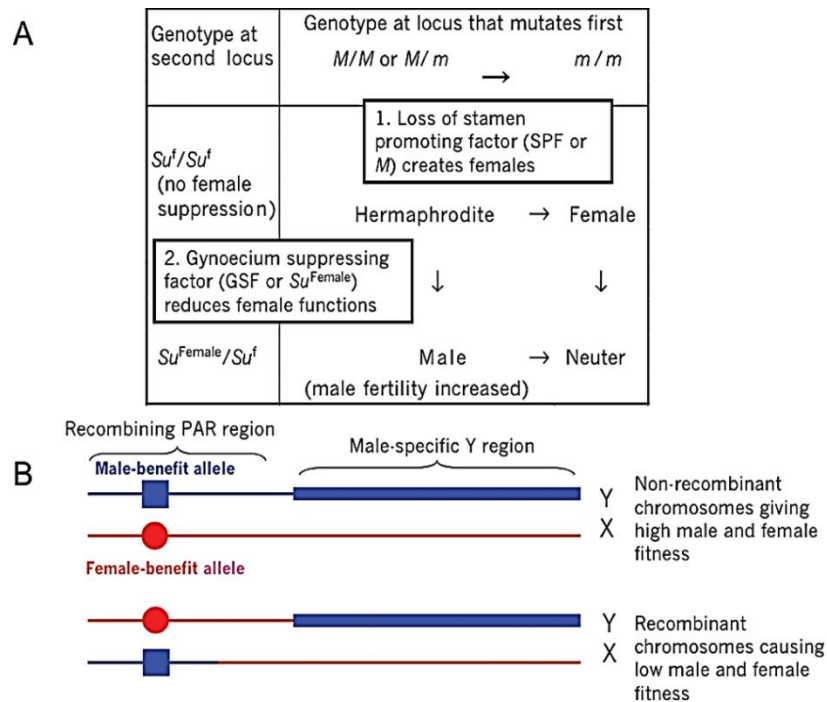
Figure 1-3. **Diagram showing situations during the evolution of dioecy.** Two mutations will segregate (i.e. although the mutations can increase in the population, they do not spread throughout all individuals and become fixed in the population), and lead to selection for reduced recombination. (A) Two sterility mutations causing female and male phenotypes are prevented from going to fixation because each causes sterility of the other sex. (B) A sexually antagonistic mutation that benefits males, but lowers the fitness of females, may also establish a polymorphism in the population at a locus in the recombining region closely linked to the male-determining region of the Y-chromosome. Reproduced from Charlesworth (2013).

Another evolutionary route via androdioecy is relatively rare (Charlesworth and Guttman, 1999). Female-sterility mutations ($Su^f \rightarrow Su^{Female}$) can occur resulting in male individuals among self-pollinating hermaphrodites (Figure 1-3). But unlike male-sterile plants, these female-sterile (male) plants would not provide any outcrossing advantage because it is assumed that the hermaphrodites can self-fertilize most of the ovules which are no longer available for female-sterile plants to fertilize. Hence, the genetic contribution of female-sterile plants to the next generation is very low compared to that of the selfing hermaphrodites. However, in case of outcrossing hermaphrodites, female-sterile plants can spread in the population if their pollen production and fertility is twice or more compared to the hermaphrodites.

Evolution of dioecy from distyly (individuals polymorphic for anther and style positions as two different bisexual floral types) is another route. Species showing distyly often also show self-incompatibility. One floral type exhibits short style and long stamens to promote pollen dispersal, and another floral type exhibits long style and short stamens. During evolution of dioecy from distyly, plants with short style flowers evolve into male and vice-versa (Lloyd, 1979).

1.3 Sex chromosomes in plants: Young and still evolving

Even though sex chromosomes in plants (Blackburn, 1923, Kihara and Ono, 1925) were identified approximately at the same time as in animals (Painter, 1921), there is very little knowledge about their structure and sequences compared to sex chromosomes in humans, mouse, fruit fly and worm. It has been almost a century since the discovery of sex chromosomes in *Rumex acetosa* (Kihara and Ono, 1925) and *Silene dioica* (Blackburn, 1923) and yet the presence of heteromorphic sex chromosomes is well established only in 19 of the dioecious species, while 20 of the dioecious species are known to have homomorphic sex chromosomes (Ming *et al.*, 2011). This indicates how remarkably rare the sex chromosomes in the plant kingdom are! Plant chromosomes are intriguing because they provide a unique opportunity for the study of sex chromosome evolution. Sex chromosomes in plants are very young and have evolved very recently (Moore *et al.*, 2003, Navajas-Pérez *et al.*, 2005, Volz and Renner, 2008, Yu *et al.*, 2008) (Table 1-1).

Table 1-1. Age of sex chromosomes in some well-studied organisms. Plant sex chromosomes are much young compared to animal sex chromosomes which makes them suitable candidates for study of sex chromosome evolution.

Species	Age of sex chromosomes
Plant Species	
Silene latifolia	$\approx$ 10 million years
Carica papaya	0.5-2.2 million years
Rumer acetosa	15-16 million years
Bryonia dioica	< 10 million years
Animal Species	
Humans	$\approx$ 150 million years
Drosophila and *C. elegans*	>100 million years

Taking into account the recent genetic and genomic studies on the male-specific regions in sex chromosomes of plant species, Ming *et al.* (2011) modified the three stage model by Westergaard (1958) to a refined six stage model for the evolution of sex chromosomes (Figure 1-4). **Stage 1** is represented by the appearance of male- and female- sterility mutations in close proximity on a chromosome. **Stage 2** involves suppression of recombination between the sex-determining loci and close flanking regions. This results in formation of a small male-specific region on Y (MSY). **Stage 3** involves expansion of MSY (male-specific region on Y) through accumulation of retrotransposons, translocations and duplications. Many of the Y-linked genes degenerate leading to lethality of the YY genotype in this stage and further stages of sex chromosome evolution. X and Y appear to be homomorphic, but at molecular level they are heteromorphic. **Stage 4** involves further expansion of MSY region and sex chromosomes appear to be heteromorphic from this stage. At this stage, Y-chromosome can be significantly larger compared to the X-chromosome, due to significant accumulation of DNA through retrotransposons, translocations and duplications. **Stage 5** represents severe degeneration and loss of non-functioning regions of Y-chromosome, eventually leading to its shrinkage. Recombination between X and Y occurs only at the terminal pseudoautosomal regions (PAR). **Stage 6** shows in further shrinkage of the Y-chromosome and complete loss of recombining pseudoautosomal regions. This eventually results in loss of complete Y-chromosome and evolution of a new system for determination of sex based on the ratio of X:Autosome.
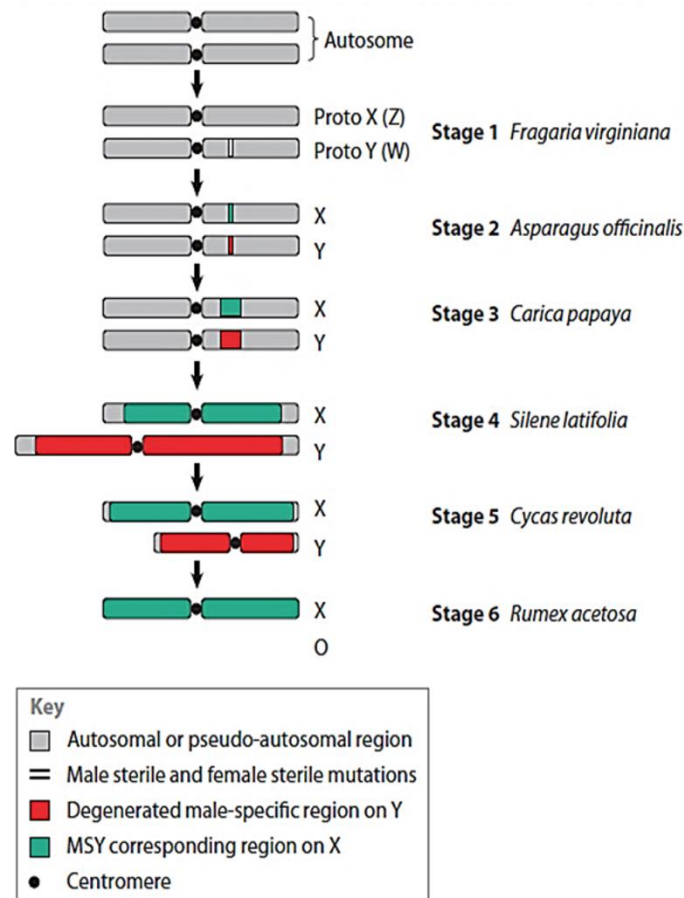
Figure 1-4. **The six-stage model for sex chromosome evolution.** Stage 1: Unisexual mutation of two sex determination genes with complementary dominance. Stage 2: Suppression of recombination, YY genotype is viable. Stage 3: Small male-specific region of the Y-chromosome (MSY) region evolved, YY genotype is not viable. Stage 4: MSY expansion and degeneration on Y-chromosome. The X and Y-chromosomes become heteromorphic. Stage 5: Deletions and severe degeneration on Y-chromosome. Stage 6: Complete suppression of recombination. Eventual loss of Y-chromosome and evolution of X: Autosome ratio mechanism for sex determination. Reproduced from Ming *et al.* (2011).

The plant species with most well characterized sex chromosomes include *Silene latifolia* (Westergaard, 1946, Matsunaga *et al.*, 1996, Guttman and Charlesworth, 1998, Delichère *et al.*, 1999, Filatov *et al.*, 2000, Scutt *et al.*, 2002), *Rumex acetosa* (Parker and Clark, 1991, Rejón *et al.*, 1994, Shibata *et al.*, 1999, Lengerova and Vyskot, 2001) and *Carica papaya* (Liu *et al.*, 2004, Ma *et al.*, 2004). *Silene latifolia* is the most well-studied dioecious plant which shows "active Y mechanism" of sex determination. The heteromorphic sex chromosomes are well-studied in *S. latifolia* (Caryophyllaceae), in which male and female plants carry XY and XX sex chromosomes respectively (Westergaard, 1940) (Figure 1-5A). *Coccinia grandis*, a dioecious climber that belong to Cucurbitaceae family also features heterogametic male (XY) and

homogametic females (XX), with Y being two times larger than the X-chromosome (Kumar and Viseveshwaraiah, 1952, Sousa *et al.*, 2013) (Figure 1-5B). The Y-chromosome is reported to be the largest of all chromosomes in *S. latifolia* (Matsunaga *et al.*, 1994). Another well-studied dioecious plant is *Rumex acetosa*. It shows X to autosome ratio as a mechanism of sex determination (Ainsworth *et al.*, 1995). The chromosome constitution of female is 12 + XX (X-chromosome: Autosome = 1) and of male is $12 + XY_1Y_2$ (X:A = 0.5), where both Y_1 and Y_2 chromosomes are required for the normal development of pollens (Parker and Clark, 1991) (Figure 1-5C). Homomorphic yet degenerate Y-chromosome is known to be present in male plants of dioecious plant *Carica papaya* (Ming et al., 2011) (Figure 1-5D). Interestingly, sex determination in *Populus* occurs through a ZW system in which the female is the heterogametic gender (Yin *et al.*, 2008). Thus, sex chromosome studies from different dioecious plant species can provide an insight for better understanding of sex chromosome evolution.

Silene latifolia   *Coccinia grandis*   *Rumex acetosa*   *Carica papaya*
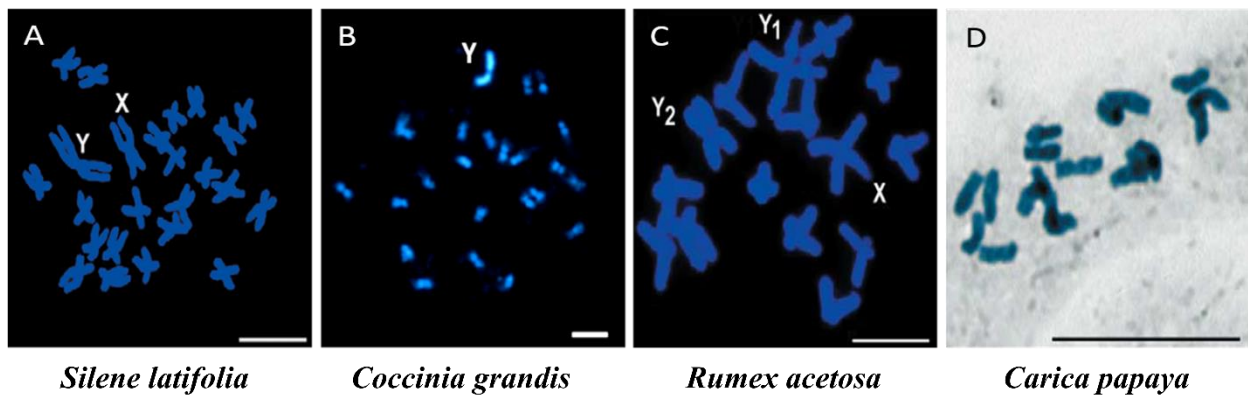
Figure 1-5. **Mitotic metaphase plates of some dioecious plants.** (A) *Silene latifolia*, (B) *Coccinia grandis* and (C) *Rumex acetosa* showing the presence of heteromorphic X and Y sex chromosomes. (D) *Carica papaya* sex chromosomes cannot be distinguished since they are homomorphic. Adapted from Sousa *et al.* (2013), Vyskot and Hobza (2004).

1.4 Molecular basis of dioecy

It is assumed that in most cases, dioecy has evolved from hermaphroditism. This means that the dioecious species have to evolve some mechanism for suppression of inappropriate sex organs (stamens in case of female plant and carpels in case of male plant) in order to develop a unisexual flower. Indeed, in *S. latifolia,* male flowers show the presence of rudimentary gynoecium (filament) and female flowers show the presence of rudimentary stamens (staminodes) (Ainsworth *et al.*, 1997a) (Figure 1-6).
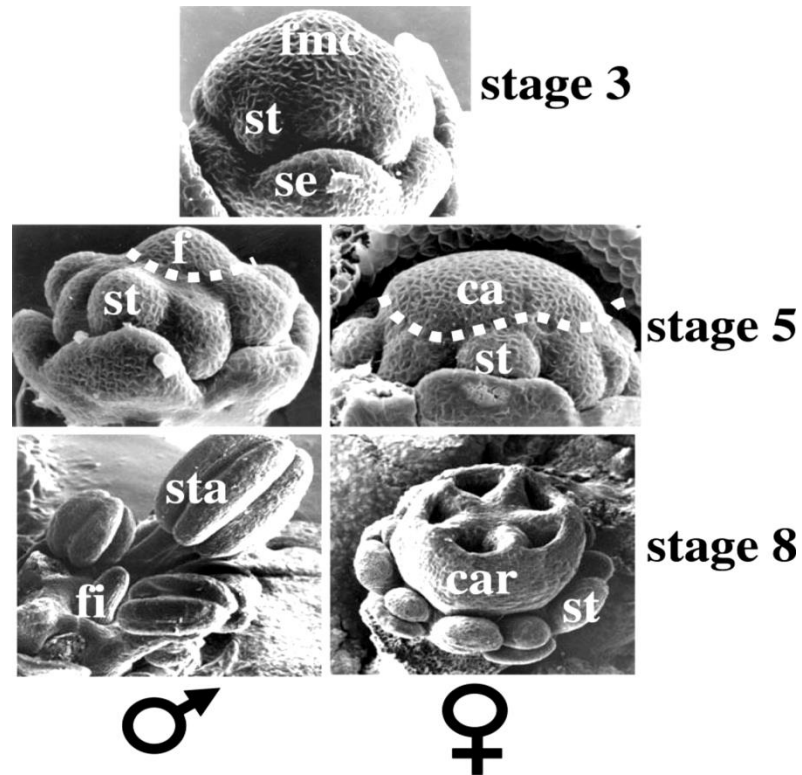
Figure 1-6. **Arrest of inappropriate sex organs in *Silene latifolia*.** At stage 3, sepal (se) and stamen (st) primordial are visible, male and female buds look identical. At stage 5, all primordial are visible; female bud can be easily distinguished by observing the wide central dome of carpel primordia (ca). At stage 5, in male bud, the central dome is much smaller and corresponds to filament primordia (f). At stage 8, five fused carpels (car) are visible in the female bud while in male bud, there is just an undifferentiated filament (fi) at the centre. At stage8, developing stamens (sta) are visible in male bud while in female, stamen primordia (st) remain arrested. se, sepal primordia; st, stamen primordia; fmc, flower meristem centre; ca, carpel primordia; f, filament primordia; sta, stamen; fi, filament; car, carpel. Reproduced from Zluvova *et al.* (2006).

From the ABC model of floral development, it is well established that certain MADS box genes grouped as Organ Identity Genes (OIGs), control the floral organ identity and development (Causier *et al.*, 2010). Of these genes, B-class genes are known to play a role in the development of petals and stamens, while C-class genes are shown to be involved in the development of stamens and carpels. Hence, these genes are considered as obvious targets for the search of sex-determining genes in plants with unisexual flowers. OIGs have been identified from numerous dioecious species including *Silene latifolia* (Hardenack *et al.*, 1994), *Rumex acetosa* (Ainsworth *et al.*, 1995), *Spinacia olereceae* (Pfent *et al.*, 2005), *Asparagus officinalis* (Park *et al.*, 2003), *Thalictrum dioicum* (Di Stilio *et al.*, 2005) and others. B- and C-class genes of *Rumex acetosa* showed very distinct expression patterns compared to hermaphrodite plants with a C-class gene

RAP1 showing sex-specific expression. Expression of *RAP1* ceases as soon as the development of carpel primordia gets arrested in male flowers and the development of stamen primordia gets arrested in female flowers. In spinach, using knock-down approach by viral induced gene silencing, Sather *et al.* (2010) showed that sexual dimorphism occurs through the regulation of B-class floral organ identity gene (Figure 1-7). However, in *Silene latifolia* it was observed that the expression of *MADS* box genes was very much similar as predicted from hermaphrodites showing slight differences between male and females. Thus, it was concluded that MADS box genes do not play a significant role in sex determination in *S. latifolia*.
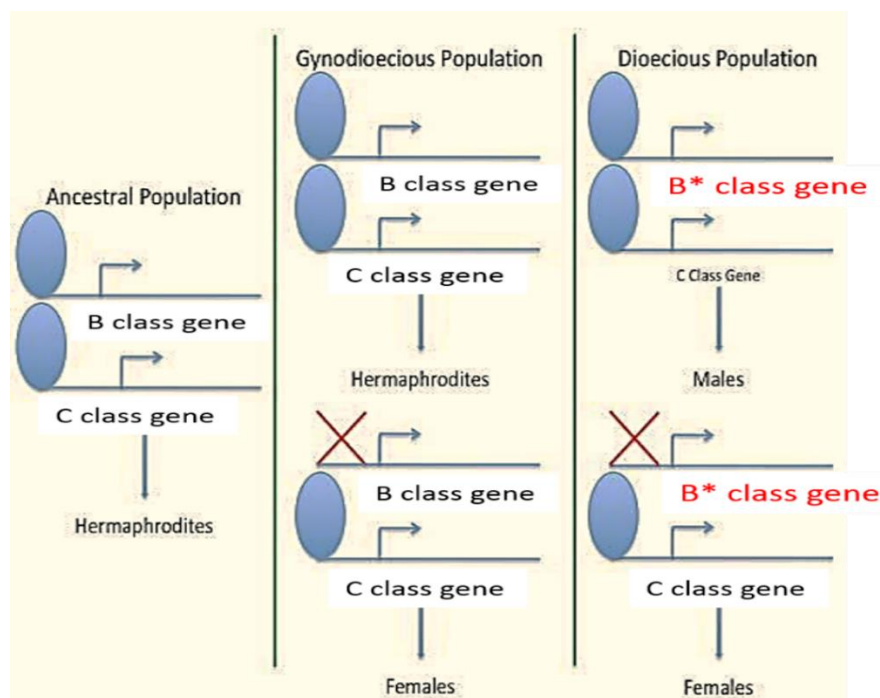
Figure 1-7. **Model for the evolution of dioecy in spinach.** In the ancestral hermaphroditic species, upstream elements, including but not limited to GA and LFY, activate both B (*PI* and *AP3*) and C (*AG*) class genes. Both classes of genes retain organ identity functions as described in the ABC model. Mutations in the B class genes, notated by *, result in premature termination of the flower in the third whorl, and thus the loss of the carpel. The resultant flower is male. Inactivation or suppression of expression of the B class genes, modulated by the GA response pathway, results in the expression of *AG* only. The absence of B class gene products causes a reduction in the number of organs in the first whorl and the formation of a single, terminal carpel. The resultant flower is female. Reproduced from Sather *et al.* (2010).

Cytological studies of Y-chromosome deletions in *S. latifolia* has led to the identification of three functionally different regions on Y-chromosome (Figure 1-7). Westergaard (1946) examined three Y deletion mutants, two of which were hermaphroditic with fully mature and

functional stamens as well as carpels. Both of these deletions were on non-recombining p-arm of Y-chromosome and overlapping with each other indicating the presence of at least one gynoecium/female suppressing locus (Su^F) (Figure 1-8). Another Y deletion mutant studied by Westergaard was asexual. The fourth whorl comprising of carpels was similar to wild-type male flowers wherein both had arrested carpels. The asexual mutant showed growth of stamens, but the pollens remained immature. This asexual mutant was partially female fertile, which helped Westergaard to map the mutation on the q arm of the Y-chromosome (Figure 1-8). This region is known as 'late male fertility restorer' (Lebel-Hardenack *et al.*, 2002). Later, Farbos *et al.* (1999) studied one more kind of asexual mutants. This mutant showed arrest of stamens at very early stage of development like in female flowers of *S. latifolia*. However, carpel was also arrested at rudimentary stage. This mutant had deletions in the p arm of the Y-chromosome between 'gynoecium suppressor' and 'late male fertility restorer' loci (Figure 1-8). This locus is named as Stamen Promoting Factor (SPF). However, till date, there is no information about the genes present in these regions that are actually responsible for the observed mutant phenotypes.
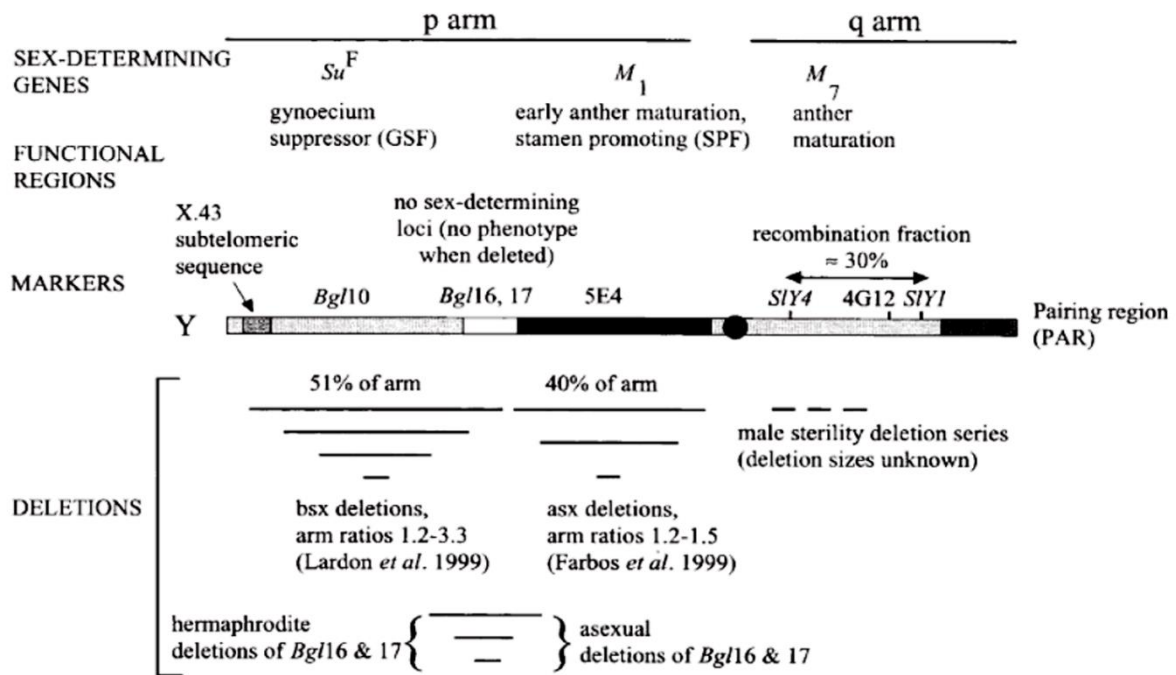
Figure 1-8. **The *Silene latifolia* Y-chromosome, showing identified markers and genes.** The deletions causing hermaphroditism (bsx), early-stage anther arrest (asx) and pollen sterility or late-stage anther arrest are shown (male sterility deletion series). Based on the approximate locations of these deletions, three sex-determining regions are proposed: stamen promoting factor (SPF), gynoecium suppressing factor (GSF) and late fertility restorer / anther maturation region. Reproduced from Charlesworth (2002).

Sex determination in papaya is simple involving a single locus M with three alleles (M_1, male; M_2, hermaphrodite and m, female). Males have heterozygous M_1m genotype, and hermaphrodites have heterozygous M_2m genotype. Females have homozygous recessive mm genotype. Combinations of homozygous dominant alleles M_1M_2, M_1M_1, and M_2M_2 showed embryonic lethality (Storey, 1953). Males are heterogametic, and sex determination follows XY mechanism. X and Y-chromosomes of papaya are homomorphic, but the Y-chromosome is highly degenerated. Non-recombining MSY region shows comparatively much high accumulation of retroelements, 85.6% compared to 52% in rest of the genome (Ming *et al.*, 2008).

1.5 Molecular basis of monoecy

Sex determination in monoecious species has been well studied in the members of Cucurbitaceae. Two unlinked loci were hypothesized in governing the sex expression of melon namely *andromonoecious* (*a*) and *gynoecious* (*g*). The genotype of different sexual forms of melon is *A-G-* (monoecious), *aaG-* (andromonoecious), *AAgg* (gynoecious) and aagg (hermaphrodite). Recently by Dr. Bendahmane's group, IPS2, France have identified the *CmACS-7* genes in *A* locus. It encodes 1-aminocyclopropane-1-carboxylic acid synthase (ACS), the rate-limiting enzyme in ethylene biosynthesis (Boualem *et al.*, 2008). *In situ* hybridization studies have shown that *CmACS-7* expresses specifically in carpel primordia. Mutation in some of the residues of ACS leads to stamen development. In another report, the same group showed that the *G* locus encodes *CmWIP1* transcription factor (Martin *et al.*, 2009). The recessive *g* allele from gynoecious and hermaphrodite plants has an insertion of a DNA transposon *Gyno-hAT* at 3'end, 1.3 kb from the stop codon. *Gyno-h*AT insertion resulted in initiation and maintenance of DNA methylation on *CmWIP1* gene leading to a loss of function, which promoted carpel development (Martin *et al.*, 2009). These two genes functionally interact with each other giving rise to a variety of sexual form in melon (Figure 1-9A). *CmACS-7* is expressed in carpel primordia to suppress stamen development, and CmWIP1 epistatically controls *CmACS-7* by suppressing carpel development. Hence, the buds expressing *CmWIP1* will exhibit carpel inhibition resulting in lack of *CmACS-7* expression that suppresses stamen development, leading to the formation of male flowers. Loss of function mutation in *CmWIP1* allows carpel development where *CmACS-7* is expressed that leads to the stamen inhibition and hence the

development of female flowers. Loss of function mutation in both *CmWIP1* and *CmACS-7* leads to the development of hermaphrodite flowers in melon (Figure 1-9A). Recently, they also discovered that the expression of *CmWIP1* is negatively regulated by *CmACS11* (Boualem et al., 2015). Hence, the expression of *CmACS11* will lead to female flower, whereas its absence will lead to a male flower (Figure 1-9A). Similar sex-determining mechanisms involving the role of ACS enzyme are also reported in other cucurbits such as cucumber and watermelon (Boualem *et al.*, 2009, Boualem *et al.*, 2016).

In another monoecious plant, maize, many mutants affecting sex expression have been characterized. Out of the six *tasselseeds* mutants affecting male flower (tassel) development, *ts4* encodes a microRNA *zma-miR172e* that targets *indeterminate spikelet1* (*ids1*) which encodes floral homeotic AP2 transcription factor (Chuck *et al.*, 2007). The dominant *Ts6* mutant has G to T transversion in *ids1* gene near the binding site of *ts4*. This prevents the silencing of *ids1* by *zma-miR172e.* Because of the failure to silence *ids1,* both *ts4* and *Ts6* mutants fail to suppress female organs in male flowers (Figure 1-9A).
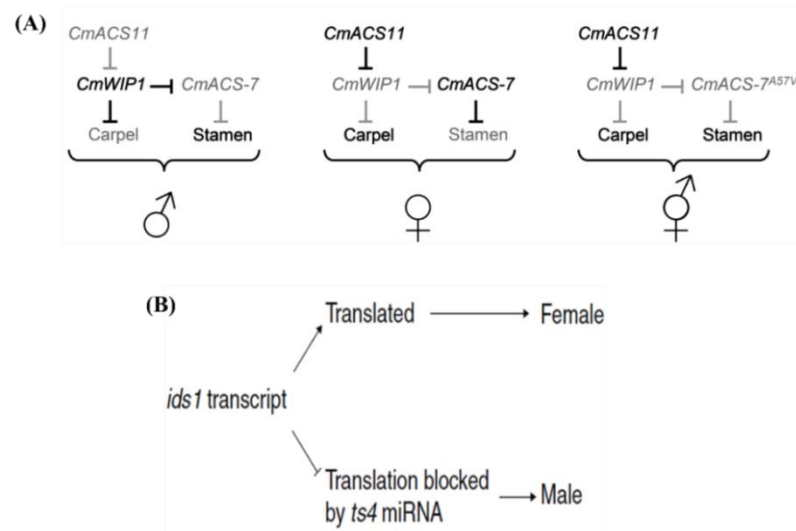
Figure 1-9. **Molecular mechanisms of sex determination in monoecious melon and maize.** (A) Genetic interaction between *CmACS11*, *CmWIP1*, and *CmACS-7* genes governs the sex in melon. Expression of *CmACS11* leads to the formation of female flower. Lack of *CmACS11* expression results in male flowers whereas *CmACS11* expression in combination with mutation in *CmACS-7* leads to the formation of hermaphrodite flowers. (B) The sex of the tassel florets depends on the translation of the *IDS1* transcript. In normal male florets, *IDS1* mRNA is not translated because of the presence of the *ts4* miRNA (miRNA172). If an *IDS1* protein is produced, either as a result of a loss-of-function mutation in the *ts4* gene or a mutation in the miRNA-binding site in *ids1* (which has occurred in the mutant *Ts6*), the floret is female. Reproduced from Boualem *et al.* (2015), Banks (2008).

1.6 Modification of sex expression in dioecy

The molecular mechanisms evolved by dioecious plants for determination of sex are not very robust. There are various reports demonstrating the modification of sex expression in dioecious plants. Janoušek *et al.* (1996) showed that 5-azacytidine (5-azaC) mediated hypomethylation in otherwise male plants of *S. latifolia* resulted in the development of hermaphrodite flowers showing various degrees of gynoecium development and fruit set (Figure 1-10). Hypomethylated plant behaved as andromonoecious bearing male and hermaphrodite flowers (Figure 1-10A). This trait was transmitted to offspring for several generations when hermaphrodite flower was self-crossed or used as pollen donor (Figure 1-10D). This clearly indicates that epigenetic mechanisms also play an important role in sex determination. Female flowers didn't show any apparent phenotype upon 5-azacytidine treatment. Also, Law *et al.* (2002) showed that silver compounds ($AgNO_3$ and $Ag_2S_2O_3$) were able to bring about enhancement of stamen growth in female plants (XX genotype) of *Silene latifolia*.
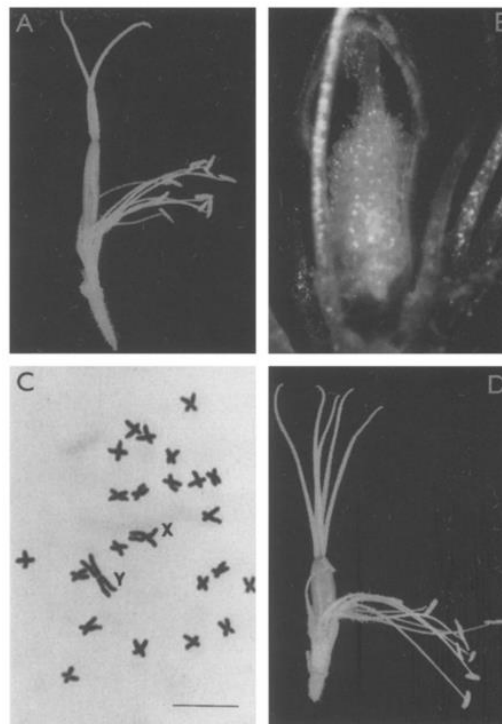
Figure 1-10. **Characteristics of androhermaphrodite plants obtained after 5-azaC treatment.** (A) Bisexual flower with a reduced ovary and two styli only. (B) Ovary (front wall removed) with fertile ovules from a nearly perfectly bisexual flower. (C) Representative male karyotype of the A1 androhermaphrodite (sex chromosomes indicated, scale bar 10 µm). (D) Perfect hermaphrodite flower from the self-progeny of A1. Reproduced from Janoušek *et al.* (1996)

However, the ethylene inhibition by silver compounds was not found to be involved in the sex modification. Pollens formed on the stamens in female flower were also not fertile. Silver compounds exhibited no effect on the development of stamens in male plants. Another unique case of sex modification found in nature is smut fungus (*Microbotryum violaceum)* infection. This smut fungus was reported to induce development of anthers in female flowers (XX genotype) of S*ilene latifolia* (Uchida et al., 2003). However, in this case, also the pollens were sterile indicating that presence of Y-chromosome is essential for the fertility of pollens.

These labile sex expressions indicate that the sex determination mechanisms are not very rigid (Korpelainen, 1998). Other opinion is that lability in sex expression might provide scope for flexibility towards environment giving advantage of better adaptation (Charnov and Bull, 1977). Our observation with regard to $AgNO_3$-mediated sex modification in *Coccinia grandis* has been described in the section 1.8.4.

1.7 Modification of sex expression in monoecy

Plant growth regulators are known to have tremendous effect on the sex expression in many monoecious plants belonging to the Cucurbitaceae family. Laibach and Kribben (1949) showed for the first time that sex of a plant can be altered by exogenous application of chemicals. They showed that exogenous application of auxin on cucumber plants can alter the sex expression towards femaleness. Another report showed that foliar application of gibberellic acid promoted male flowers in monoecious cultivars of cucumber, where normally female flowers would have developed (Wittwer and Bukovac, 1958) Also, application of silver compounds such as $AgNO_3$ and $Ag_2S_2O_3$ is shown to masculinize monoecious plants of cucumber (Yin and Quinn, 1995). $AgNO_3$ is well-known inhibitor ethylene response (Beyer, 1976). Hence, ethylene might also be playing a role in sex determination of monoecious cucurbits. In accordance with these, application of ethrel (an ethylene releasing compound) promoted the development of female flowers at higher frequency in monoecious melon, cucumber and summer squash genotypes (Stankovic *et al.*, 2005).

Since we were interested in establishing *Coccinia grandis* as a model system to understand dioecy, we began with detailed morphological, histological and cytological studies with three different sexual forms of *C. grandis* (Ghadge *et al.*, 2014). Present author has actively

contributed as a co-author in this publication and has continued research on this topic for doctoral research. The detail findings are described in the following sections.

1.8 *Coccinia grandis* (Cucurbitaceae)- a potential model system to study dioecy

Coccinia grandis is a dioecious plant belonging to the family Cucurbitaceae, a family well-known for its diverse sexual systems. All species belonging to genus *Coccinia* are dioecious. *Coccinia grandis* has received less attention as a system to understand dioecy compared to *S. latifolia*, *R. acetosa*, *Populus*, etc. *C. grandis*, commonly known as ivy gourd, is used as a vegetable and has rich medicinal value. It has male and female flowers on separate unisexual plants. Male plants possess 22 autosomes and XY sex chromosomes; whereas females have 22 autosomes and XX sex chromosomes. Sex chromosomes of *C. grandis* are heteromorphic with male-specific Y-chromosome being at least twice the size of the biggest autosome. Apart from male and female plants, we have also identified a rare gynomonoecious form (GyM) during our initial survey of *C. grandis* population in the northeastern part of India (Tripura). Gynomonoecious plant bears two types of flowers: morphologically hermaphrodite (GyM-H) and pistillate (GyM-F). Phylogenetic analysis based on the sequences of matK and trnSGCU-trnGUCC intergenic spacer regions clearly established the identity of gynomonoecious form as another sexual form of *C. grandis*. Also, we have demonstrated that a foliar spray of AgNO$_3$ on female plants of *C. grandis* induces the development of morphologically hermaphrodite flowers by eliminating the stamen inhibition.

1.8.1 Morphological differences amongst three sexual phenotypes

Among the flowers of male, female and gynomonoecious *C. grandis*, sepal and petal characteristics are found to be strikingly similar. All the flower types have five sepals and five fused petals. However, morphological differences exist among them with regards to the androecium and gynoecium whorls. Male flowers have only three whorls with five (2+2+1) synandrous stamens and no sign of carpels (Figure 1-11A, E). In contrast, female flowers have all the four whorls with rudimentary stamens and gynoecium consisting of three carpels having fused style with three bifid stigmas (Figure 1-11B, F). Morphologically, the pistillate flowers (GyM-F) of gynomonoecious plant were identical to the flowers of female plant (Figure 1-11D, H). The hermaphrodite flowers (GyM-H) of gynomonoecious plant had four whorls with full-

sized stamens and carpels unlike pistillate flowers with arrested stamens (Figure 1-11C, G) (Ghadge et al., 2014). In few of the hermaphrodite flowers of gynomonoecious plant, we observed either petaloid stamens or incomplete growth of stamens. The ratio of hermaphrodite and pistillate flowers on the gynomonoecious plant varied with season. Also, we observed that seed content (seed number as well as seed weight) of fruits from female flowers was higher compared to that of fruits from GyM plant.
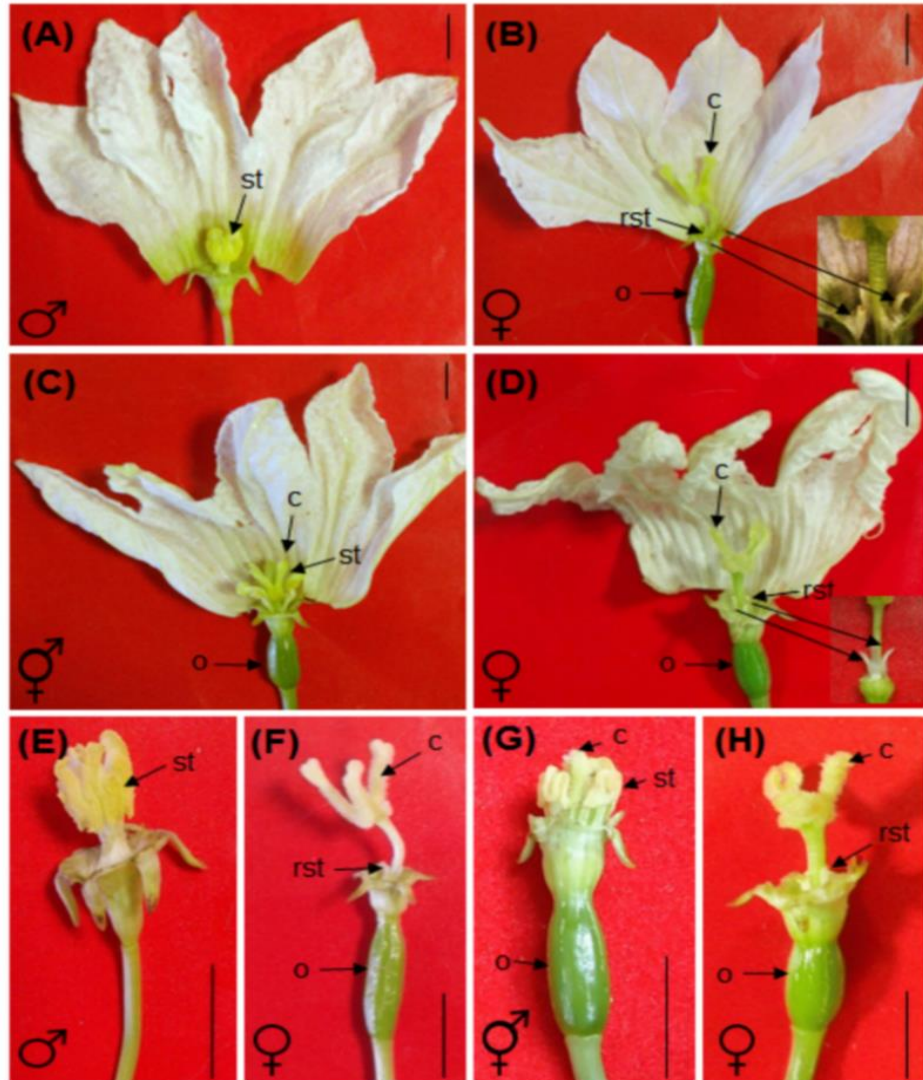
Figure 1-11. **Morphology of mature flowers of *Coccinia grandis*.** Macroscopic view of staminate flower (A) of male plant, pistillate flower (B) of female plant, hermaphrodite (GyM-H) (C) and pistillate (GyM-F) (D) flowers of gynomonoecious (GyM) plant with petals cut open. Petals removed from staminate flower (E) of male plant, pistillate flower (F) of female plant, hermaphrodite (GyM-H) (G) and pistillate (GyM-F) (H) flowers of gynomonoecious (GyM) plant to show inner floral organs. st: Stamens, c: carpels, rst: rudimentary stamens, o: ovary. Scale bars = 1 cm. Reproduced from Ghadge *et al.* (2014).

1.8.2 Histological analysis of flower buds

In order to understand the spatiotemporal development of floral organs, we have carried out histological analysis of flower buds from male, female and gynomonoecious plants at different stages of development.

Male: Longitudinal sections of male flowers at early stages of development (stages 3-5) showed the presence of sepals, petals and stamens (Figure 1-12A, B). However, carpel initials could not be observed. Even at later stages of development, the signs of carpel initials were absent. The presence of carpel primordia during the formation of floral meristem cannot be completely ruled out just based on histology. Gradual stamen growth can be observed in histological sections of middle (stages 5-7) and late (10-12) stages of flower development (Figure 1-12A-E). Pollen grains can also be observed in late stages of male buds when petals were about to open (Figure 1-12E).
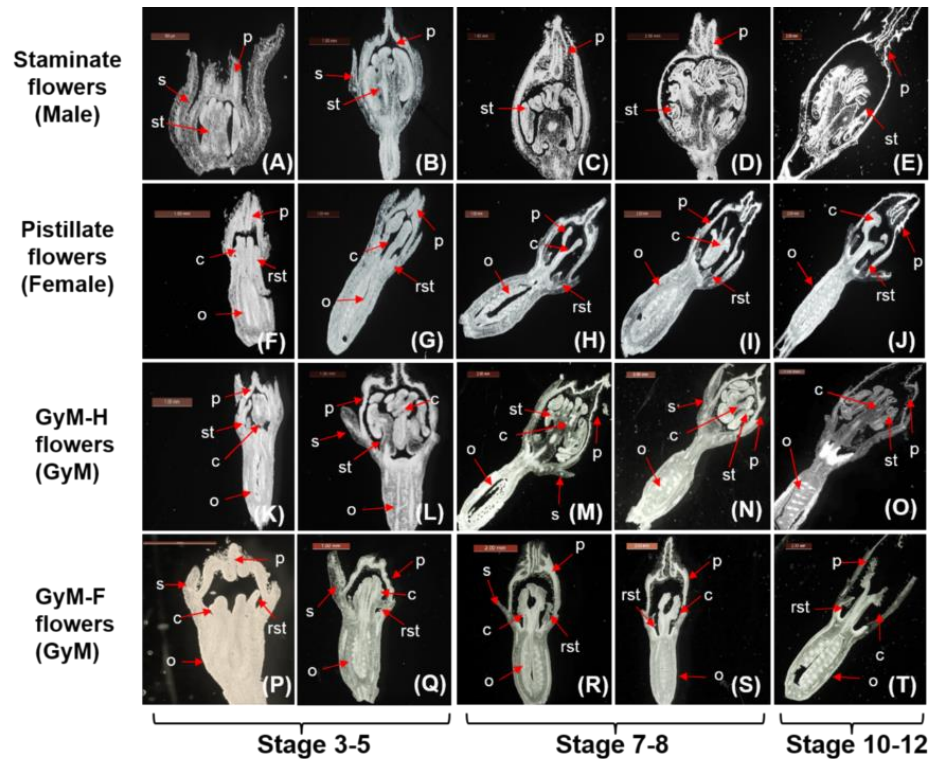
Figure 1-12. **Longitudinal sections (L.S) of *C. grandis* flower buds at different developmental stages.** (A–E) are the sections of staminate flowers of male plant, (F–J) are the sections of pistillate flowers of female plant, (K–O) and (P–T) are the sections of hermaphrodite (GyM-H) and pistillate (GyM-F) flowers of gynomonoecious (GyM) plant respectively. p: Petals, s: sepals, c: carpels, st: stamens, rst: rudimentary stamens, o: ovary. Scale bars are 500μm in A, 1mm in B, C, F, G, H, K, L, P and Q, and 2mm in D, E, I, J, M, N, O, R, S and T. Reproduced from Ghadge *et al.* (2014).

Female: At early stages of development, sepals, petals, stamens and carpels with inferior ovary can be observed in the histological sections (Figure 1-12F, G). We find that the development of androecium gets arrested in early stages (around stage 4-5) whereas, carpel growth can be observed in middle and late stages of female buds (Figure 1-12G-J).

GyM: Histological sections of GyM-F buds looked very similar to the buds of female plant (Figure 1-12P-T). However, in the GyM-H buds, all organs namely sepals, petals, stamens and carpels showed gradual development through early, middle and late stages of development (Figure 1-12K-O). Stamen arrest was observed only in the sections of GyM-F buds but not in GyM-H buds. Both GyM-H and GyM-F buds showed inferior ovary like the buds of female plant.

1.8.3 Karyotyping of the three sexual forms

Kumar and Viseveshwaraiah (1952) reported for the first time the presence of heteromorphic sex chromosomes in *C. grandis*. The male plants carry 22A+XY and female plants carry 22A+XX-chromosomes (Figure 1-13). The Y-chromosome is twice in size compared to largest autosome, while the X-chromosome is morphologically indistinguishable from the other autosomes (Figure 1-13A, B). The present record of diploid chromosome number 2n = 24 in both male and female sexes and the presence of heteromorphic sex chromosomes in male has been reconfirmed by Bhowmick *et al.* (2012), (Sousa *et al.*, 2013) and Ghadge *et al.* (2014). The gynomonoecious (GyM) form reported by Ghadge *et al.* (2014) showed 22A+XX genotype (Figure 1-13C). GyM form showed stamen development even in absence of Y-chromosome.
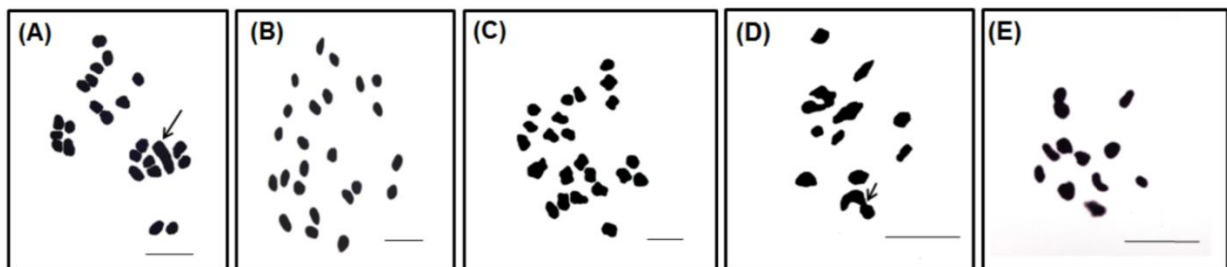
Figure 1-13. **Metaphase chromosomes of *C. grandis*.** Mitotic metaphase chromosomes showing 2*n* = 24 chromosomes of male (A) (arrow indicates the large Y-chromosome), female (B) and gynomonoecious (GyM) (C) plants. Meiotic metaphase chromosomes showing 12 bivalents of male (D) (arrow indicates end to end pairing of X and Y-chromosomes), gynomonoecious (GyM) (E) plants. Scale bar = 5 µm. Reproduced from Ghadge *et al.* (2014).

1.8.4 AgNO₃ induced sex modification

Law *et al.* (2002) showed that silver compounds (AgNO$_3$ and Ag$_2$S$_2$O$_3$) were able to stimulate stamen growth in female plants (XX genotype) of *S. latifolia*. Silver is a well-known inhibitor of ethylene signalling, and silver compounds have been shown to promote stamen growth in monoecious and dioecious cucurbits such as melon, cucumber and *Cannabis* (Sarath and Mohan Ram, 1979, Yin and Quinn, 1995). In order to see if silver nitrate (AgNO$_3$) has any effect on *C. grandis*, different concentrations of AgNO$_3$ solution were sprayed on the basal leaves of male, female and GyM plant. Newly emerging flower buds of wild-type female plants showed enhanced growth of stamens after application of AgNO$_3$ solution whereas; male flowers did not show any changes in floral structure (Ghadge *et al.*, 2014) (Figure 1-14).
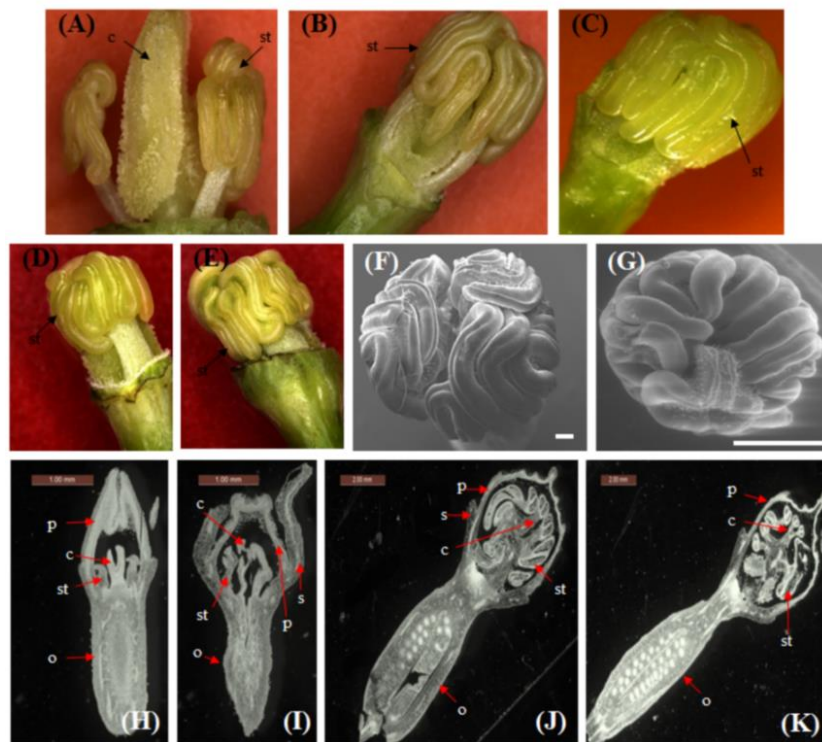
Figure 1-14. **Effects of silver nitrate (AgNO₃) solution on flowers of female *C. grandis* plant.** (A-C) are the pictures of female flowers after spraying of AgNO₃ solution showing gradual enhanced stamen growth. Magnified view of stamens in (D) pistillate flowers of AgNO₃ treated female plant and (E) hermaphrodite (GyM-H) flowers of gynomonoecious (GyM) plants. Scanning electron micrographs of top view of (F) pistillate flowers from AgNO₃ treated female plant and (G) hermaphrodite (GyM–H) flowers of gynomonoecious (GyM) plants. Petals and sepals have been removed to view sexual structures better. Longitudinal sections (H-K) of flower buds of silver nitrate treated female plant (after spraying of 35 mM AgNO₃ solution). H, I – flower buds of stage 5, J – flower bud of stage 8 and K – flower bud of stage 10. p: Petals, s: sepals, c: carpels, st: stamens, o: ovary. Scale bars are 300µm in F, 1mm in G, H and I, and 2mm in J and K. Reproduced from Ghadge *et al.* (2014).

Histological studies further confirmed the dose-dependent stamen growth in wild-type female flowers. However, $AgNO_3$ concentrations higher than 35 mM had lethal effect. At dosages of 30 and 35 mM of $AgNO_3$, the morphology of newly developed flowers was comparable to GyM-H flowers after 10-12 days of observation (Figure 1-14) (Ghadge *et al.*, 2014). Interestingly, all mature flowers in GyM plant were found to be hermaphroditic after application of $AgNO_3$, indicating that even the staminodes of pistillate flower buds have developed into mature stamens.

1.8.5 *Mating experiments and pollen fertility*

Mating experiments were carried out to assess the fertility of pollens from the flowers of male plant as well as morphologically hermaphrodite flowers (GyM-H) of gynomonoecious plant. Approximately, 90% fruit setting was observed in the cross between male and female plants. However, cross between morphologically hermaphrodite flowers (GyM-H) of gynomonoecious plant and flowers of female plant did not result in any fruit set. Similar results were achieved when flowers of female plant were replaced with pistillate flowers of gynomonoecious plant. Also, a cross made between flowers of male plant and emasculated hermaphrodite flowers (GyM-H) of gynomonoecious plant did not yield any fruit set. In another experiment, unopened morphologically hermaphrodite buds (GyM-H) of gynomonoecious plant were bagged at late stage of development to check for the possibility of self-fertilization. However, we observed that none of the bagged GyM-H flower buds set fruit. These results showed that pollens from only the flowers of male plant are fertile, whereas pollens from morphologically hermaphrodite flowers of gynomonoecious plant are sterile in nature. However, female organs of both the hermaphrodite (GyM-H) and pistillate (GyM-F) buds of gynomonoecious plant are fertile and can give rise to fruit upon fertilization with the pollens from male plant.

In order to confirm these results, we carried out pollen viability staining using acetocarmine and fluorescein diacetate (FDA). In addition to the pollens from the flowers of male plant and hermaphrodite flowers of gynomonoecious plant, pollens from the hermaphrodite flowers (Ag-H) of $AgNO_3$ treated plants were also studied. Pollens from flowers of only male plant were able to retain the acetocarmine and FDA stains (Figure 1-15A-F). While authors

observed that the pollens of GyM-H and Ag-H buds did not retain either of the stains indicating non-viable pollens.

Additionally, *in vitro* pollen germination assay was carried out from male, GyM-H and Ag-H buds. Different pollen germination media were tested by varying the sucrose concentration. Highest frequency of pollen germination (38%) was achieved when pollens of male flowers were incubated in 5% sucrose solution containing required amount of $Ca(NO_3)_2$ and H_3BO_3 (Figure 1-15G,H). In contrast, pollens of GyM-H and Ag-H buds did not show any germination when incubated in different germinating media. From the results of these three different approaches, we concluded that the pollens from only male flowers are fertile whereas pollens from GyM-H and Ag-H flowers are sterile in nature.
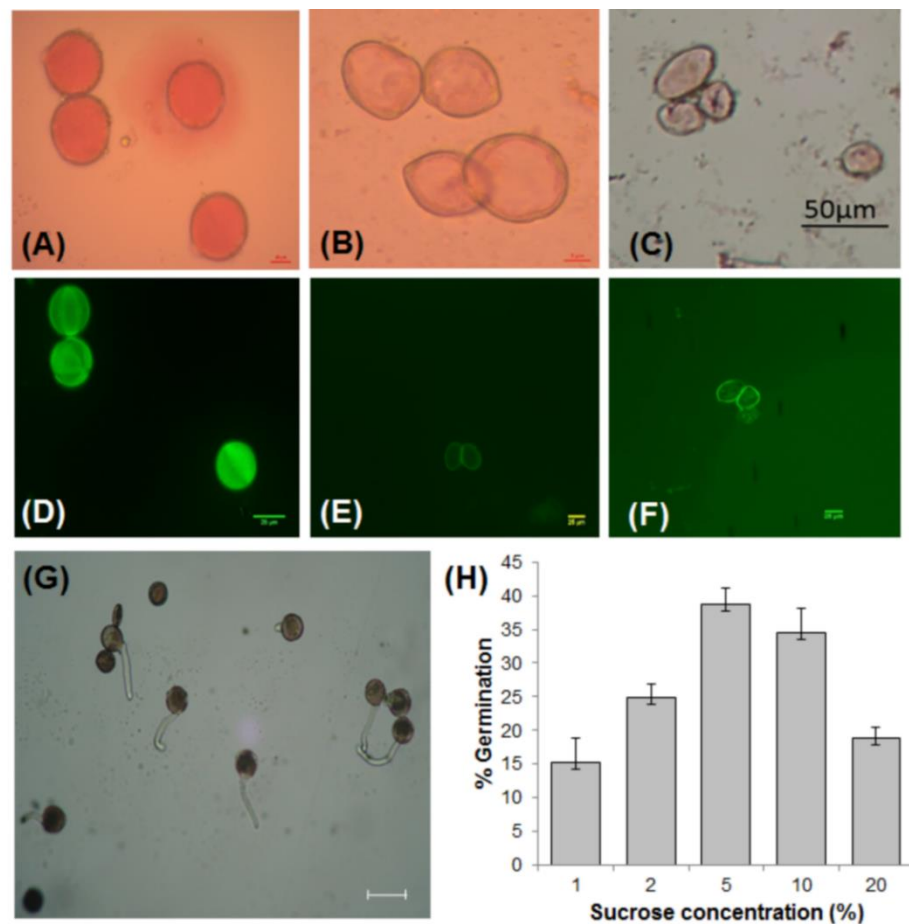
Figure 1-15. **Viability tests of pollens from male, gynomonoecious (GyM) and AgNO₃ treated female plants of *C. grandis*.** Pollens stained with 1% acetocarmine from male (A), gynomonoecious (GyM) (B) and AgNO₃ treated female (C) plants. (D), (E) and (F) are the fluorescein diacetate (FDA) stained pollens from male, gynomonoecious (GyM) and AgNO₃ treated female plants respectively. Pollens stained with acetocarmine (A) and FDA (D) are viable. Scale bars are 10 μm in A, 5 μm in B, 50 μm C, and 25 μm in

D, E and F. (G) Highest germination of male pollens in 5% sucrose solution. Scale bar = 50 μm. (H) Graphical representation of the germination percentage in different concentrations of sucrose solutions. Means ± standard errors are reported in the graph; n=10. Reproduced from Ghadge *et al.* (2014).

An interesting observation to note here is that presence of Y-chromosome is correlated with the property of pollen fertility. Only male plants have Y-chromosome, whereas both gynomonoecious and AgNO$_3$ treated female plants lack Y-chromosome. As Y-chromosome has already been shown to govern pollen development in dioecious plants like *Silene latifolia* and *Rumex acetosa*, we hypothesize that Y-chromosome might have a similar role in governing pollen fertility of *C. grandis*.

Our morphological, histological and cytological characterization, uncovered a few interesting events during the unisexual flower development in *C. grandis*, such as inhibition of stamens at early stages of development in flowers of female plant but not in the male and GyM-H flowers. However, if AgNO$_3$ treatment was given on the leaves at optimal concentration, the flowers of female plants were able to bypass this inhibition indicating a probable role of ethylene in stamen suppression. Also, we observed a correlation between the presence of Y-chromosome and fertility of pollens. Thus, in order to understand the underlying molecular mechanisms behind these interesting observations, we raised the following crucial questions with regard to sex expression and modification in *C. grandis*.

1.9 Key questions and objectives

1.9.1 Key questions

1. Which are the genes that regulate sex determination in *C. grandis*?

2. Does Y-chromosome play any role in rendering maleness?

3. How does AgNO$_3$ bring about sex modification in *C. grandis*?

4. Does ethylene play any role in sex determination? Is the stamen suppression role of *ACS* genes conserved in dioecious members of Cucurbitaceae?

In absence of genome and transcriptome data for *Coccinia,* we initiated with candidate gene approach to identify the molecular players involved in sex expression and modification.

However, since the evolution of dioecy has most likely taken place independently in different angiosperm families, the players governing sex might not be conserved. Hence, we took a *de novo* transcriptomics approach followed by proteomics in order to identify sex-biased genes and proteins in *C. grandis*. We wanted to validate the key candidate sex-biased genes by functional studies to understand their role in sex determination and differentiation. However, lack of an efficient *in vitro* regeneration and genetic transformation tool was a major limitation to study gene function in *C. grandis*. Hence, we chose to standardize virus induced gene silencing tool as an alternative strategy for gene knockdown in *C. grandis*.

To understand sex expression and modification in dioecious *C. grandis*, we laid down the following objectives for the present investigation which are described in their respective chapters:

1.9.2 Objectives

1. Candidate gene approach and comparative transcriptomics to identify novel players involved in sex expression and modification of *C. grandis*.
2. Proteomic characterization of *C. grandis* flower buds in order to understand unisexual flower development and $AgNO_3$ mediated sex modification.
3. Standardization of virus induced gene silencing (VIGS) in *C. grandis* for functional characterization of key differentially expressed genes.

Part of this chapter has been published in the following research article:

Ghadge A., Karmakar K., **Devani R.S.**, Banerjee J., Mohanasundaram B., Sinha R.K., Sinha S., Banerjee A.K. (2014). Flower development, pollen fertility and sex expression analyses of three sexual phenotypes of *Coccinia grandis*. **BMC Plant Biology.** 14: 325. DOI: 10.1186/s12870-014-0325-0

Chapter 2: Candidate gene approach and comparative transcriptomics to identify novel players involved in sex expression and modification

2.1 Background

Monoecy, dioecy, and hermaphroditism are the three major sexual forms observed among the flowering plants. Ninety (90 %) of angiosperms are found to be hermaphrodite (both male and female organs are in the same flower), while 5 % plant species exhibit monoecy (male and female flowers are on the same plant) and remaining 5 % show dioecy (male and female flowers are in separate plant) (Charlesworth, 2002). Dioecism provides a unique opportunity to study the genetic basis of sex determination. *Silene latifolia* (Caryophyllaceae), *Rumex acetosa* (Polygonaceae), *Carica papaya* (Caricaceae), *Spinacia oleracea* (Chenopodiaceae) and *Populus* (Salicaceae), have been well characterized to understand the mechanism of sex determination (Ainsworth *et al.*, 1995, Pfent *et al.*, 2005, Yin *et al.*, 2008, Urasaki *et al.*, 2012). However, the molecular mechanism and the genes that govern sex determination are not well understood.

Coccinia grandis (L.) Voigt, a dioecious member of Cucurbitaceae family having an inferior ovary has received comparatively less attention. Members of Cucurbitaceae family exhibit variety of sexual forms (Kouonon *et al.*, 2009). Apart from its rich medicinal value, C *grandis*, commonly known as ivy gourd, is also used as a vegetable. *Coccinia grandis* bears male and female unisexual flowers on separate plants. Similar to *Silene latifolia* (Caryophyllaceae), the sex in *Coccinia grandis* is determined by the presence of Y-chromosome (Kumar and Deodikar, 1940, Bhaduri and Bose, 1947, Chakravorti, 1948). The chromosome constitution of male and female plants is 22A + XY and 22A + XX respectively, where Y-chromosome is larger than the X-chromosome (Kumar and Viseveshwaraiah, 1952, Bhowmick *et al.*, 2012, Sousa *et al.*, 2013). The male flower consists of three convoluted (bithecous) stamens (Bhuskute *et al.*, 1986, Ghadge *et al.*, 2014) and lacks female reproductive organs; however, the female flower consists of three rudimentary stamens surrounding the three fused carpels with an inferior ovary (Ghadge *et al.*, 2014). There are two ways by which unisexual flower development can be achieved. One of the ways is when both male and female sex organ primordia are initiated at early stages of flower development, but at later stages, the opposite sex organs are aborted as in *Silene latifolia* (Hardenack *et al.*, 1994). Another way is that organ primordia of the opposite sex organs do not develop at all as shown in *Thalictrum dioicum* (Di Stilio *et al.*, 2005). Also, there are flowers, wherein the inappropriate sex organs are retained in rudimentary form (instead of getting aborted) as in *Rumex* and *C. grandis* (Ainsworth *et al.*, 1995, Ghadge *et al.*, 2014).

Additionally, *Coccinia grandis* shows sex modification upon application of AgNO$_3$ leading to the development of stamens in female flower (Such flower will be referred to as Ag-H) as described in our previous report (Ghadge *et al.*, 2014). Ag$^+$ has been long known to be an inhibitor of ethylene response (Beyer, 1976). It has been suggested that the binding of Ag$^+$ to the ethylene receptor inhibits the conformational change, which maintains the receptor in the active conformation (McDaniel and Binder, 2012). Application of silver compounds such as silver nitrate (AgNO$_3$) or silver thiosulphate (Ag$_2$S$_2$O$_3$) masculinizes monoecious plants such as *Cucumis sativus* as well as female plants of dioecious species such as *Silene latifolia* and *Cannabis sativa* (Sarath and Mohan Ram, 1979, Yin and Quinn, 1995, Law *et al.*, 2002). However, the mechanism of action by which Ag$^+$ induces stamen development is not known till date (Law *et al.*, 2002).

Despite the interesting discovery of sex chromosomes in dioecious plants more than 50 years ago, the mechanism of sex determination remains poorly understood (Blackburn, 1923, Kihara and Ono, 1925). This is primarily because of the slow pace at which sex-linked genes were identified from dioecious species (one to two genes/year) (Muyle *et al.*, 2012). However, the improvement in NGS technology has already started changing the situation by accelerating the rate of sex-linked gene identification. The NGS-based approach has a big advantage that it does not require prior knowledge of the gene sequences to be investigated. Recently, an NGS-based RNA-Seq approach was applied to *Silene latifolia,* which was the first report that demonstrated the phenomenon of dosage compensation in plants (Muyle *et al.*, 2012). A comparative transcriptomics approach was applied to papaya, a trioecious species, to identify the candidate genes for sex determination. This study led to the identification of 312 unique tags that were specifically mapped to the primitive sex chromosome (X or Y^h) sequences in papaya (Urasaki *et al.*, 2012). A genome-wide transcriptional profiling of apical tissue of a gynoecious mutant (Csg-G) and the monoecious wild-type (Csg-M) of cucumber was also performed to isolate genes involved in sex determination. This study revealed that genes involved in plant hormone signalling pathways, such as *ACS*, *Asr1*, *CsIAA2*, *CS-AUX1*, and *TLP*, and their crosstalk might play a critical role in the sex determination. Authors have also predicted the regulation of some transcription factors, including *EREBP*-9, in sex determination (Wu *et al.*, 2010). In another study, transcriptome sequencing was carried out from cucumber flower buds of two near-isogenic lines, WI1983G, a gynoecious plant which bears only pistillate flowers and

WI1983H, a hermaphroditic plant which bears only bisexual flowers (Guo *et al.*, 2010). This study identified differentially expressed genes as well as putative SSR and SNP markers between flowers of two different sexes. Akagi *et al.* (2014) sequenced genomic DNA, mRNA as well as small RNA from flower buds of persimmon and identified a Y-chromosome–encoded small RNA, *OGI,* that targets a homeodomain transcription factor *MeGI* regulating pollen fertility in a dosage-dependent manner. A recent *de novo* transcriptomics study in garden *Asparagus* identified 570 differentially expressed genes, where genes involved in pollen microspore and tapetum development were shown to be specifically expressed in males and supermales in contrast to females (Harkess *et al.*, 2015).

Initially, we began with candidate gene approach and isolated the *C. grandis* homologs of sex-determining genes that have been studied in other species. We chose organ identity genes, *ACC synthase* (*ACS*) and *miR172* as potential candidates based on literature review. Organ identity genes have been shown to inhibit female organs in male flowers of spinach and *Thalictrum* (Di Stilio et al., 2005, Pfent et al., 2005, Sather et al., 2010). In maize, a monoecious system, *miR172-APETALA 2* module inhibits the development of female organs in male flower (Chuck *et al.*, 2007). Hence, we also studied the expression of *miR172* in sex forms of *C. grandis*. Sex determination in monoecious relatives of *C. grandis* such as melon, cucumber and watermelon is governed by *ACS* (Boualem *et al.*, 2008, Boualem *et al.*, 2009, Boualem *et al.*, 2016). Also, the role of *CmACS11* has been implicated in evolution of dioecious systems from monoecious cucurbit melon (Boualem *et al.*, 2015). Hence, *ACS* was an obvious choice for study of sex determination in *C. grandis*. Expression analysis of these candidate genes was undertaken to understand their role in unisexual flower development of *C. grandis*. However, the role of sex-determining genes from one model system may not be conserved in other systems. Hence, a *de novo* transcriptomics approach was also taken to identify novel sex-biased genes in *C. grandis*.

In order to elucidate the mechanism of sex expression and AgNO$_3$ mediated sex modification in *C. grandis*, a comprehensive RNA-Seq study from early-staged male (M), female (F), GyM-H and Ag-H as well as middle-staged male and GyM-H flower buds was carried out. *De novo* transcriptome was assembled to identify *C. grandis* homologs of various flower development genes. Digital expression profiling was undertaken to identify sex-biased genes which might play a pivotal role in the arrest of stamen development in female flowers,

genes that promote anther development in female flowers upon AgNO$_3$ treatment and genes controlling pollen fertility in male flowers.

2.2 Methods

2.2.1 *Flower bud collection and RNA isolation*

Clones of wild-type male, female and gynomonoecious (GyM) forms of *C. grandis* were grown in the experimental plot at IISER Pune, India. Gynomonoecious (GyM) plant bears pistillate flowers (GyM-F) as well as morphologically hermaphrodite flowers (GyM-H) (Herbarium Voucher: Tripura University Campus, Karmakar, 433). Foliar spray of 35 mM AgNO$_3$ solution to the basal leaves on some of the female plants led to the development of morphologically hermaphrodite flowers (Ag-H) as per our earlier observation (Ghadge *et al.*, 2014). Ag-H flower buds were morphologically similar to GyM-H flowers. Flower buds from male (M), female (F), GyM-H and Ag-H were harvested separately in liquid nitrogen and categorized into early and middle stages based on our previous study (Ghadge *et al.*, 2014) (Figure 2-1).
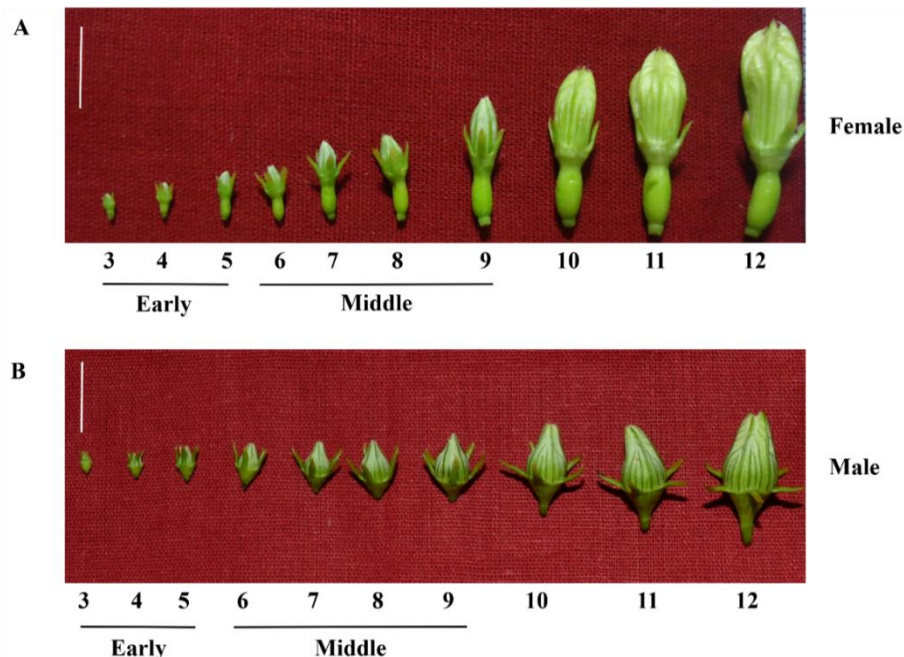
Figure 2-1. **Different stages of *C. grandis* flower buds selected for RNA-Seq analysis.** (A) Female, (B) Male flower buds. GyM-H and Ag-H buds sized similar to female buds were chosen for RNA seq study as described in our previous work (Ghadge *et al.*, 2014). Scale bar = 1 cm. Reproduced from Devani *et al.* (2017).

In early-staged male flower buds, only stamens are present with no sign of carpel initials. Whereas, early-staged female flowers (stages 3-4, Figure 2-1) have both carpel and stamen primordia. Stamen growth in female flowers gets arrested around stages 4–5. In the hermaphrodite flowers of gynomonoecious plant, however, both stamens and carpels develop simultaneously during early as well as middle stages of development. Our selection of early-staged flower buds was carried out such that the event of stamen inhibition in female flowers can be analysed. Whereas, middle-staged flower buds were chosen such that meiosis-stage and pollen maturation event can be investigated. Total RNA was isolated by TRIzol reagent (Invitrogen) following the manufacturer's instructions. RNA quality was assessed using an Agilent Bioanalyzer RNA nanochip, and RNA samples with RIN > 8.0 were used for library preparation.

2.2.2 Identification and isolation of PISTILLATA, AGAMOUS, miR172, and ACC Synthase homologs from C. grandis

To isolate *AGAMOUS* (*CgAG*) and *PISTILLATA* (*CgPI*) homologs from *C. grandis*, degenerate primers were designed from conserved sequences of PI and AG using iCODEHOP (Boyce *et al.*, 2009). Two µg of total RNA pooled from male, female and GyM-H buds was used for RT-PCR reactions using SuperScript® III One-Step RT-PCR System with Platinum®Taq (Invitrogen). The first step of reaction included incubation at 50 °C for 20 min for cDNA synthesis followed by 94 °C for 2 min, 40 cycles of incubations at 94 °C for 15 s, 50 °C for 30 s and 68 °C for 35 s. Final extension was carried out at 68 °C for 5 min. Amplified products were resolved on 2 % agarose gel and cloned into pGEMT Easy vector (Promega) and finally sequence verified. These sequences were used to design primers for 5′ and 3′ RACE to obtain the full-length transcript sequences. RACE-ready cDNAs were generated using SMARTer RACE cDNA synthesis kit (Clontech). 5′ and 3′ sequences were further amplified from the cDNAs using the designed primers and the universal primer provided with the kit. Amplified 5′ and 3′ regions of *CgPI* and *CgAG* were sequence verified. Primers were designed to amplify full-length transcripts. Deduced amino acid sequences were aligned with cucumber and melon homologs of PISTILLATA and AGAMOUS using Clustal Omega and consensus sequences were shaded using Boxshade server. Conserved domains were identified using NCBI's Conserved Domain Database (CDD) search. For isolation of *ACC synthase* (*ACS*) homolog of *C. grandis,* primers were designed based on the

DNA sequence alignment of *CmACS-7* and *CsACS2* using PriFi (Fredslund *et al.*, 2005). Two micrograms total RNA was used for complementary DNA (cDNA) synthesis by SuperScript III reverse transcriptase (Invitrogen) using an oligo(dT) primer. Touchdown PCR was carried out using Advantage II polymerase with following steps: 94 °C for 2 min, 5 cycles of incubation at 94 °C for 30 s, 63 °C for 30 s and 72 °C for 150 s. Annealing temperature was lowered to 61 °C for the next 5 cycles and 59 °C for the following 5 cycles keeping the rest of cycling conditions the same. Finally, 30 cycles of incubation at 94 °C for 30 s, 57 °C for 30 s and 72 °C for 150 s were carried out followed by final extension at 72 °C for 5 min. Amplified product was resolved on 1.5 % agarose gel and cloned into pGEMT Easy vector (Promega) and finally sequence verified. Mature *miR172* was detected by stem-loop RT-PCR as described earlier (Varkonyi-Gasic *et al.*, 2007). Pulsed reverse transcription was carried out using stem-loop primer miR172STP. End-point PCR was performed using miR172FP and universal reverse primer (univRP). The amplicon was cloned in the pGEM-T Easy (Promega) and was confirmed by sequencing.

2.2.3 RNA-Seq library preparation and sequencing

Library preparation was performed at Genotypic Technology's Genomics facility, Bangalore using Illumina TruSeq RNA Sample Preparation Kit according to the manufacturer's specifications. RNA sequencing libraries were prepared in duplicate for early-staged male (M), female (F), GyM-H and Ag-H flower buds, as well as middle-staged male (M) and GyM-H flower buds. The quality of all the twelve libraries and insert size distribution was assessed using an Agilent High Sensitivity Bioanalyzer Chip. Libraries showed a peak in the range of 250-1000 bp. The effective sequencing insert size was 130-880 bp, and the inserts were flanked by adaptors whose combined size was 120 bp. Libraries were quantified using Qubit and sequenced on Illumina NextSeq 500 platform, producing 2 X 150-nucleotide paired-end reads. RNA-Seq data generated in this study has been deposited in the NCBI SRA study SRP111347.

2.2.4 Pre-processing of Illumina reads and de novo transcriptome assembly

Raw RNA-Seq reads were processed using Trimmomatic v0.33 for trimming adapters as well as low-quality bases from ends of the reads (Bolger *et al.*, 2014). Poor quality reads with average Phred quality score < 20 and reads with length < 36 were also filtered out. The resulting set of good quality reads were then assembled with Trinity v2.1.1 software using default

parameters (Grabherr *et al.*, 2011, Haas *et al.*, 2013). Similarly, *de novo* transcriptome was also assembled using another software, Bridger with *k-mer* length = 25 (Chang *et al.*, 2015).

The quality of the resulting assembly was assessed by various methods. BUSCO (Benchmarking Universal Single-Copy Orthologs) was used to explore completeness of the transcriptome according to conserved ortholog content (Simão *et al.*, 2015). Finally, TransRate was used to compare the assembly to the publicly available *Cucumis sativus* protein-coding primary transcript sequences (Smith-Unna *et al.*, 2016). The assembly quality was comparable between the two software (Trinity and Bridger) when assessed by BUSCO and Transrate.   The downstream analyses is much more streamlined for Trinity assembly as the software package comes bundled with the scripts required for transcript quantification, differential expression analysis and functional annotation, etc . Also, there is an active community for technical support in order to help with troubleshooting the problems that one might encounter while using Trinity. Hence, Trinity assembly was chosen for all the downstream analyses reported in this study. RNA-Seq read representation of the assembly was checked using bowtie2 (Langmead and Salzberg, 2012). Ex90N50 transcript contig length (the contig N50 value based on the set of transcripts representing 90 % of the expression data) was computed using *contig_ExN50_statistic.pl* script bundled with Trinity. Then the representation of full-length reconstructed protein-coding genes was studied. The assembled transcripts were compared with SwissProt using BLAST and the hits were analyzed using a perl script *blast_outfmt6_group_ segments.tophit_coverage.pl,* provided with the Trinity package.

2.2.5 Annotation of the de novo-assembled transcripts

The *de novo*-assembled transcripts were compared with the viridiplantae sequences from nr and SwissProt database using BLASTX with an e-value threshold of 1e-3 (Altschul *et al.*, 1990). BLAST output generated from this comparison was loaded into BLAST2GO for mapping GO terms to the transcripts and annotation (Götz *et al.*, 2008). Enzyme codes and KEGG pathway mapping were also carried out. ANNEX (Annotation Expander) was used to enhance the annotations. Finally, GO-Slim mapping was applied to get a broad overview of the ontology content.

In addition to BLAST2GO facilitated annotation, Trinotate pipeline was used to carry out comprehensive functional annotation of the transcripts leveraging various annotation databases

(eggNOG/GO/KEGG databases) (Haas *et al.*, 2013). Trinotate pipeline also included identification of open reading frames, homology search against SwissProt and TrEMBL. Protein domain identification was carried out using HMMER/PFAM. Protein signal peptide and transmembrane domains were predicted by signalP and tmHMM respectively.

2.2.6 Transcript quantification and differential expression analysis

align_and_estimate_abundance.pl script from Trinity package was applied to align cleaned reads from each library to the *de novo* transcriptome using bowtie and to estimate the transcript abundance using RSEM (Li and Dewey, 2011). *abundance_estimates_to_matrix.pl* script was used to construct a matrix of counts and a matrix of normalized expression values. *PtR* script was used to generate correlation matrix and Principal Component Analysis (PCA) plot for comparing replicates across all the samples. Differential expression analysis was carried out with two biological replicates from the count matrix using *run_DE_analysis.pl* with edgeR as the method of choice (Robinson *et al.*, 2010). *analyze_diff_expr.pl* script was used to examine GO enrichment and to extract all transcripts that had p-values at most 1e-3 and were at least 2^2 fold differentially expressed. The DE features were partitioned into clusters with similar expression patterns by *define_clusters_by_cutting_tree.pl* script with Ptree method.

2.2.7 Validation of differentially expressed genes by qRT-PCR

For expression analysis, qRT-PCR was carried out using aliquots of the same RNA samples that were used for RNA sequencing. Two micrograms (2 µg) of total RNA was used for complementary DNA (cDNA) synthesis by SuperScript IV reverse transcriptase (Invitrogen) using an oligo(dT) primer. *CgACT2* gene was used as reference gene for normalization. qRT-PCR was performed on BIO-RAD CFX96 machine with gene-specific forward and reverse primers. The reactions were carried out using Takara SYBR Premix Ex Taq II (Takara Bio Inc.) and incubated at 95 °C for 3 min followed by 40 cycles of 95 °C for 15 s, 58 °C for 15 s and 72 °C for 15 s. PCR specificity was checked by melting curve analysis, and data were analysed using the $2^{-\Delta\Delta CT}$ method.

2.2.8 *Primers used in this study*

Primer name	Sequence (5´ -> 3´)
CgPI B12 F	GGAAAAAGACTGTGGGATGCNAARCAYGA
CgPI D8 R	TCTTTCTTGCAGATTTGGTTGNATNGGYTG
CgAG A26 F	GAGGAAAGATTGAAATTAAGAGAATHGARAAYAC
CgAG A51 R	CTCTCAGCTTAGCAGCTTCYTGYTGRTA
PI GSP1 (5' RACE)	TGGTTGGATCGGTTGCACTCTGAAGGC
PI GSP2 (3' RACE)	TCACTGGTGTTCGTGAGAAGCAGTCGGAG
AG GSP1 (3' RACE)	TCTATGTGATGCTGAAGTTGCTCTAATCG
AG GSP2 (5'RACE)	TAGTGGAATCTGAGGATGCCTTCTTGTATC
CgPI RTB-F	GGGAAGAGGCTGTGGGATGCTAAG
CgPI RTB-R	CTCCGACTGCTTCTCACGAACACC
CgAG RTC-F	GAGGCCAGATTGGAGAGAGGAATCAG
CgAG RTC-R	GTTTGTTGTTGTTGCTGCTGCTGTTG
18S-FP2	GGGCATTCGTATTTCATAGTCAGAG
18S-R	CGGTTCTTGATTAATGAAAACATCCT
CgACSpf F	AGATTCGCCGTATTTTGCTGGCTGGAAAGCGTATG
CgACSpf R	AGCGTAAAGCTAGACATCCGGCGAGCGG
CsACS_Ex-Intr2_Fwd	CCTTACTATCCTGGATTTGACAGAG
CsACS_7252_CtoG	AATGTCTTCGATTGTGGACCGTTG
172STP	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATGCAG
172FP	CGGCGGTAGAATCTTGATGATG
miRNA Universal RP	GTGCAGGGTCCGAGGT
DEX1_qF	CTATGTCTGTTCCCCTGCCCTG
DEX1_qR	GACCAGACTTTTCAATCACCCATG
CER3_qF	TCCCAACACTTAGACTCTCAACAC
CER3_qR	GGAAAGGTTCACTATTGGGCGTG
EMS1_qF	TCCGTCAAGTTGTTACACACCAC
EMS1_qR	TGCTCCACTCCTCTGTCTCTCC
TPD1_qF	TGCAACTGAAATGACACTCGCC

TPD1_qR	ACCAATCTAACAGGCCACACAC
ZAT3_qF	ACGACGGAAACTAAGCGGCG
ZAT3_qR	GCGCCGAAGATCACTCCTCC
DYT1_qF	GAAGCGGAAAGGGGCATTCAC
DYT1_qR	TGACGTGCAAAAGAGCCCATTG
AMS_qF2	GGGAGCTCTGGATTGTCTTCGGG
AMS_qR2	TCTCCGGCTTCCTTTATCCCGC
MMD1_qF2	TTCAGTGTGATCAATTGTGTCTCGC
MMD1_qR2	TTGGAAGGTGGGGCTACAGATTC
MS1_qF2	CGACCGAGCAACAGGTGAAAC
MS1_qR2	TCCTTCAGCTCATCAATGGTGGC
FERONIA_qF	AGATCCACCTCCCACTACACCC
FERONIA_qR	ACCGCAACTCCAACCACAACAG
EIL1_qF	GAGAACGCTAACTTTCCACGCC
EIL1_qR	TGGAAACGAGGAATGGTGGCC
SHT_qF	CAAGGTTGTGGCTGCCGATG
SHT_qR	GTTGTCGTCGTTGTTGCGGG

2.3 Results

2.3.1 Candidate gene isolation and expression analysis

In order to understand whether B and C class Organ Identity Genes (OIGs) have any role in determining the sex of the developing flowers of male, female and GyM plant, *CgPI* (a B class OIG) and *CgAG* (a C class OIG) were isolated and an expression analysis was carried out using quantitative real-time PCR (qRT-PCR). The degenerate primers based on the conserved amino acid sequences of PI (PISTILLATA) and AG (AGAMOUS), yielded ~350 bp of *PISTILLATA* (*CgPI*) and ~250 bp of *AGAMOUS* (*CgAG*) homologs through RT-PCR reaction. The partial sequences for *CgPI* [DDBJ: AB859715] and *CgAG* [DDBJ: AB859714] have been deposited in DDBJ. Full-length transcript sequences were deduced from 5′ and 3′ RACE products, and amplicons of *CgPI* (~893 bp) and *CgAG* (~952 bp) were obtained (Figure 2-2A).
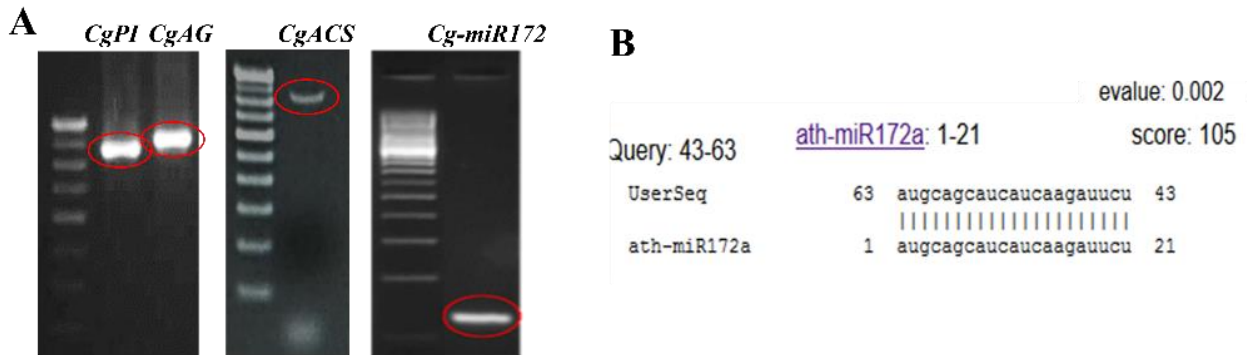
Figure 2-2. **Isolation of candidate genes from *C. grandis*.** (A) Amplification of full-length *CgPI*, *CgAG*, *CgACS* and *Cg-miR172* transcripts from total RNA harvested from flower buds. Adapted from (Ghadge *et al.*, 2014) (B) Sequence confirmation of *Cg-miR172* using miRBase search.

cDNA for *CgPI* and *CgAG* coded for putative proteins of 212 and 232 amino acids respectively. The deduced amino acids sequences for both the genes showed high conservation when aligned with PISTILLATA and AGAMOUS sequences from melon and cucumber (Figure 2-3A, B). Two consensus regions, MADS domain and K-box, were found on the deduced amino acid sequences (Figure 2-3A, B). *CgPI*, a B class gene required for petal and stamen development, was found to be expressed in male, wild-type female and GyM flower buds (Figure 2-4A). Expression of *CgAG*, a C class gene essential for stamen and carpel development, was also noted in male, wild-type female and GyM flower buds (Figure 2-4B). Our results showed that both these genes are expressed in all developmental stages (early, middle and late) of flowers from male, female and GyM plant. *CgPI* had a significant difference of expression across all three sexual forms during early, middle and late developmental stages (Figure 2-4A), while *CgAG* showed significant differential expression in buds of early stages only (Figure 2-4B). We have also noted that *CgPI* expression is comparatively high in male flower buds than that of wild-type female buds. However, GyM flowers exhibited an intermediate level of *CgPI* expression in early and late staged buds (Figure 2-4A).

A

```
At_PI      1   --------------------------------------MGRGKIEIKRIEN
Cg_PI      1   --------------------------------------MGRGKIEIKRIEN
Cs_CUM26   1   --------------------------------------MGRGKIEIKRIEN
Cm_PMADS2  1   MNTKTQYRMGPLYVHYFKNPYVCFEELKRKEEEKRKKMGRGKIEIKRIEN

At_PI     14   ANNRVVTFSKRRNGLVKKAKEITVLCDAKVALIIFASNGKMIDYCCPSMD
Cg_PI     14   SSNRQVTYSKRRNGIIKKAKEITVLCDAQVSLIIFASSGKMHEYCSPSTP
Cs_CUM26  14   SSNRQVTYSKRRNGIIKKAKEITVLCDAQVSLVIFASSGKMHEYCSPSTP
Cm_PMADS2 51   SSNRQVTYSKRRNGIIKKAKEITVLCDAQVSLVIFASSGKMHEYCSPSTP

At_PI     64   LGAMLDQYQRLSGKKLWDAKHENLSNEIDRIKKENDSLQLELRHLKGEDI
Cg_PI     64   LIDILDKYHKQSGKRLWDAKHENLSNEMDRVKKENDNMQIELRHLRGEDI
Cs_CUM26  64   LVDILDKYHKQSGKRLWDAKHENLSNEMDRVKKENDNMQIELRHLRGEDI
Cm_PMADS2 101  LVDILDKYHKQSGKRLWDAKHENLSNEMDRVKKENDNMQIELRHLRGEDI

At_PI     114  QSLNLKNLMAVEHAIEHGLDKVRDHQMEILISKRRNEKMMAEEQRQLTFQ
Cg_PI     114  TSLNYKELMALEEALENGLTGVREKQSEFMKMMRTNERMMEEENKRLNYE
Cs_CUM26  114  TSLNYKELMALEEALENGLTGVREKQSEFMKMMRTNERMMEEENKRLNYE
Cm_PMADS2 151  TSLNYKELMSLEEALENGLTGVREKQSEFMKMMRTNERMMEEENKRLNYE

At_PI     164  LQQQEMAIASNARGM-------MMR--DHDGQFGYRVQPIQPNLQEKIMS
Cg_PI     164  LYQKEMAAMGDSVREMDNGYNQRMRDFNSQMPFAFRVQPIQPNLQERIN*
Cs_CUM26  164  LYQKEMVAMGDSVREMDIGYNQRMRDFNSQMPFAFRVQPIQPNLQERE--
Cm_PMADS2 201  LYQKEMVAMGDSVREMDIGYNQRMRDFNSQMPFAFRVQPIQPNPQERE--

At_PI     205  LVID
Cg_PI     213  ----
Cs_CUM26       ----
Cm_PMADS2      ----
```

B

```
CgAG          1   ----------------MGRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYE
Cs_MADS1      1   ----------------MGRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYE
Cm_AGAMOUS    1   -------------------------------------------------
Mc_MADS_box2  1   --------MSDSPQRKMGRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYE
At_AGAMOUS    1   MAYQSELGGDSSPLRKSGRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYE

CgAG         35   LSVLCDAEVALIVFSTRGRLYEYANNSVRGTIERYKKAFADSINSGLSVA
Cs_MADS1     35   LSVLCDAEVALIVFSTRGRLYEYANNSVRGTIERYKKAFADSSNSGLSVA
Cm_AGAMOUS    1   -------------------SFKIGSVRGTIERYKKAFADSSNSGLSVA
Mc_MADS_box2 43   LSVLCDAEVALIVFSSRGRLYEYANNSVKATIDRYKASSDSSNTG-STS
At_AGAMOUS   51   LSVLCDAEVALIVFSSRGRLYEYSNNSVKGTIERYKKAISDNSNTG-SVA

CgAG         85   EANVQFYQQEASKLKRQIREIQNSNRHILGEALSSLSIKDLKSLEARLER
Cs_MADS1     85   EANVQFYQQEATKLKRQIREIQNSNRHILGEALSSLPKELKSLEGRLER
Cm_AGAMOUS   30   EANVQFYQQEATKLKRQIREIQNSNRHILGEALSSLPKELKSLEGRLER
Mc_MADS_box2 92   EANTQFYQQEAAKLRVQIGNIQNSNRNMLGESLSSLSVKDLKSLESKLEK
At_AGAMOUS  100   EINAQYYQQEASKLRQQIISIQNSNRQLMGETIGSMSPKELRNLEGRLER

CgAG        135   GISKVRAKKNETLCAEMEFMQKREMELQNHNNYLRAQIAEHERIQQQQ-Q
Cs_MADS1    135   GISKVRAKKNETLFAEMEFMQKREMELQSHNNYLRAQIAEHERIQQQQQQ
Cm_AGAMOUS   80   GISKVRAKKNETLFAEMEFMQKREVELQSHNNYLRAQIAEHERIQQQQQQ
Mc_MADS_box2 142  GISRIRSKKNETLFAEIEYMRKREIDLHNNQLLRAKIAESERNASMIG-
At_AGAMOUS  150   SITRIRSKKNEILFSEIDYMQKREVDLHNDNQILRAKIAENERNNPSISL

CgAG        184   QQQTNMMQRATYESVG----GQYDENR--SYGAVG-LMDSDNHYA---HQD
Cs_MADS1    185   QQQTNMMQRATYESVG----GQYDDENRSTYGAVGALMDSDSHYA--PQD
Cm_AGAMOUS  130   QQQTNMMQRATYESVG----GQYDDENRSTYGAVGALMDSDSHYA--PQD
Mc_MADS_box2 191  ---------GDRE------LMQSHPYDPRDFFQVNGLQHNNHQY--PRQ
At_AGAMOUS  200   MP-----GGSNYPQLMPPPQTQSQPFDSRNYFQVAALQPNNHHYSSAGRQ

CgAG        225   HLTALQLV*--
Cs_MADS1    229   HLTALQLV---
Cm_AGAMOUS  174   HLTALQVEVTL
Mc_MADS_box2 224  DNMALQLV---
At_AGAMOUS  245   DQTALQLV---
```

Chapter 2: *De novo* transcriptomics approach for *C. grandis*

```
C      CgACSpf          1  ------------------------DSPYFAGWKGYDEDPYNESTNPSGVIQMGLAENQ
       CsACS2           1  MAIEIDIEQNSSVELSRIGTSETHGEDSPYFAGWKAYDEDPYNESTNPSGVIQMGLAENQ
       CmACS-7_Vedrant  1  MAIEIDIEQNPTVELSRIGTSETHGEDSPYFAGWKAYDEDPYNESTNPSGVIQMGLVENQ
       CmACS-7_PI12411  1  MAIEIDIEQNPTVELSRIGTSETHGEDSPYFAGWKAYDEDPYNESTNPSGVIQMGLAENQ


       CgACSpf         35  VSFDLLEEHLEENYEEEAK-NCSGFRENALFQDYHGLVSFRIAMAGFMEEIRGGRAKFDP
       CsACS2          61  VSFDLLEEYLEENCEGEGNYLNSGFRENALFQDYHGLFSFRSAMGSFMEEIRGGRAKFDP
       CmACS-7_Vedrant 61  VSFDLLEEYLEENCEGEGNYLNSGFRENALFQDYHGLFSFRSAMGSFMEEIRGGRAKFDP
       CmACS-7_PI12411 61  VSFDLLEEYLEENCEGEGNYLNSGFRENALFQDYHGLFSFRSAMGSFMEEIRGGRAKFDP


       CgACSpf         94  NRVVLTAGATAANELLTFILADPGDALLVPTPYYPGFDRDLRWRTGVKIVPIHCDSSNNF
       CsACS2         121  NRVVLTAGATAANELLTFILANPGDALLVPTPYYPGFDRDLRWRTGVKIVPIHCDSSNNF
       CmACS-7_Vedrant 121 NRVVLTAGATAANELLTFILANPGDALLVPTPYYPGFDRDLRWRTGVKIVPIHCDSSNNF
       CmACS-7_PI12411 121 NRVVLTAGATAANELLTFILANPGDALLVPTPYYPGFDRDLRWRTGVKIVPIHCDSSNNF


       CgACSpf        154  QITPKALEEAYNSAMAMKIKVRGVLITNPSNPLGATIQRSTIEEILDFVTRKNIHLVSDE
       CsACS2         181  QITPKALEEAYNSATEMKIKVRGVLITNPSNPLGATIQRSTIEDILDFVTRKNIHLVSDE
       CmACS-7_Vedrant 181 QITPKALEEAYNSAMEMKIKVRGVLITNPSNPLGATIQRSTIEDILDFVTRKNIHLVSDE
       CmACS-7_PI12411 181 QITPKALEEAYNSAMEMKIKVRGVLITNPSNPLGATIQRSTIEDILDFVTRKNIHLVSDE


       CgACSpf        214  IYSGSVFSSAEFTSVAEVLESRGYKNAERVHIVYSLSKDLGLPGFRVGTIYSYNDKVVTT
       CsACS2         241  IYSGSVFSSDEFTSVAEVLESRGYKNAERVHIVYSLSKDLGLPGFRVGTIYSYNDKVVTT
       CmACS-7_Vedrant 241 IYSGSVFSSAEFTSVAEVLESRSYKNAERVHIVYSLSKDLGLPGFRIGTIYSYNDKVVTT
       CmACS-7_PI12411 241 IYSGSVFSSAEFTSVAEVLESRSYKNAERVHIVYSLSKDLGLPGFRIGTIYSYNDKVVTT


       CgACSpf        274  ARRMSSFTL--------------------------------------------------
       CsACS2         301  ARRMSSFTLISSQTQRFLASMLSNRKFTENYIKMNRDRLKKRYEMIIEGLRTAGIECLEG
       CmACS-7_Vedrant 301 ARRMSSFTLISSQTQRFLASMLSNRKFTEKYIKMNRDRLKKRYEMIIEGLRTAGIECLEG
       CmACS-7_PI12411 301 ARRMSSFTLISSQTQRFLASMLSNRKFTEKYIKMNRDRLKKRYEMIIEGLRTAGIECLEG


       CgACSpf             -----------------------------------------------------------
       CsACS2         361  NAGLFCWMNLSPLLKDKKTKEGEIEIWKRILKEVKLNISPGSSCHCSEPGWFRVCFANMS
       CmACS-7_Vedrant 361 NAGLFCWMNLSPLLKDKKTKEGEIEIWKRILKEVKLNISPGSSCHCSEPGWFRVCFANMS
       CmACS-7_PI12411 361 NAGLFCWMNLSPLLKDKKTKEGEIEIWKRILKEVKLNISPGSSCHCSEPGWFRVCFANMS


       CgACSpf             -------------------------
       CsACS2         421  EKTLHVALDRIRRFMERMKKENEAN
       CmACS-7_Vedrant 421 EKTLHVALDRIRRFMERMKKENEAN
       CmACS-7_PI12411 421 EKTLHVALDRIRRFMERMKKENEAN
```

Figure 2-3. **Multiple sequence alignment for *CgPI*, *CgAG* and *CgACS* proteins.** (A) Comparison of *CgPI* with *At_PI*, *Cg_PI*, *Cs_CUM26* and *Cm_pMADS2* (*PISTILLATA* like genes from *Arabidopsis thaliana*, *Coccinia grandis*, *Cucumis sativus* and *Cucumis melo* respectively). (B) Comparison of *CgAG* with *Cg_AG*, *Cs_MADS1*, *Cm_AGAMOUS*, *Mc_MADS_box2*, *At_AGAMOUS* (*AGAMOUS* like genes from *Coccinia grandis*, *Cucumis sativus*, *Cucumis melo*, *Momordica charantia* and *Arabidopsis thaliana* respectively). (C) Comparison of *CgACS* with *CmACS-7* and *CsACS2* from melon and cucumber. Conserved regions are shaded in black.

Further, our results for stamen-specific expression analysis showed a significant difference for both *CgPI* and *CgAG* levels between stamens of male, GyM-H, AgNO₃ treated female plant, rudimentary stamens of GyM-F and wild-type female plant (Figure 2-4C, D). Surprisingly, rudimentary stamens of GyM-F showed higher *CgPI* expression than stamens of GyM-H flowers (Figure 2-4C). Similarly, *CgACS* was amplified by touchdown PCR using primers based on the sequences of *CmACS-7* and *CsACS2*. Partial sequence of 847 bp was amplified, which showed high homology to *CgACS-7* and *CsACS2* (Figure 2-2A, Figure 2-3 C).
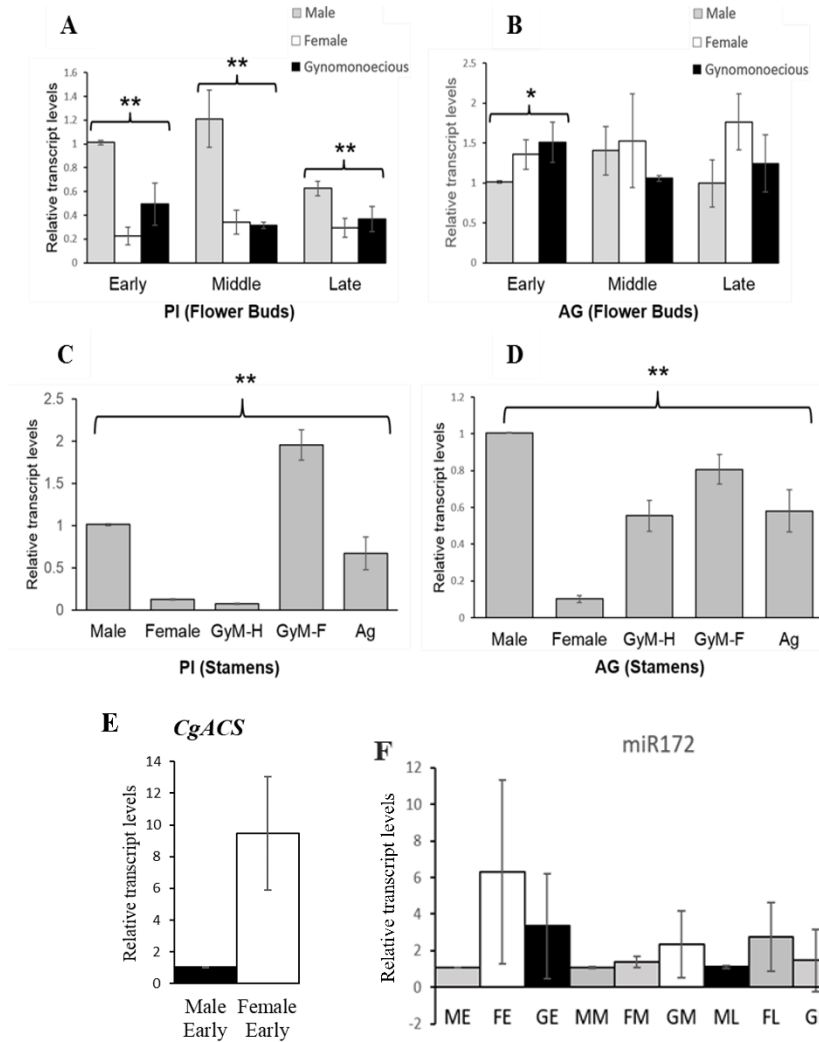
Figure 2-4. **Expression analyses of candidate genes from *C. grandis*.** Expression patterns of *CgPI* (A) and *CgAG* (B) in flower buds of male, female and gynomonoecious (GyM) *C. grandis* at different developmental stages (early, middle and late) by quantitative real-time PCR (qRT-PCR). Stamen specific expression patterns of *CgPI* (C) and *CgAG* (D) from flowers (late developmental stage) of male, female (rudimentary), hermaphrodite (GyM-H) and pistillate (GyM-F, rudimentary) flowers of gynomonoecious (GyM) and converted flowers of $AgNO_3$ treated plants. Error bars indicate SD (standard deviation) of three biological replicates each with three technical replicates. Asterisks indicate statistical differences as determined using single factor ANOVA (*$P < 0.05$ and **$P < 0.01$). Adapted from Ghadge *et al.* (2014). Expression patterns of *CgACS* (E) in flower buds of early-staged male and female buds. Expression patterns of *Cg-miR172* (F) in flower buds of male, female and gynomonoecious (GyM) *C. grandis* at different developmental stages (early, middle and late).

Primers were designed for qRT-PCR based expression analysis using the sequence of *CgACS* isolated from *C. grandis*. Interestingly, we observed that the expression of *CgACS* was approximately 10-fold higher in early-staged female buds compared to that in early-staged buds (Figure 2-4E). The sex-biased expression profile of *CgACS* was similar to the expression of

CgACS-7 and *CsACS2* in flowers of melon and cucumber respectively. This gave us an indication that role of ACS might be conserved in sex determination of *C. grandis* similar to monecious cucurbits melon, cucumber and watermelon. Also, mature *miR172* was isolated from *C. grandis* by stem-loop PCR (Figure 2-2). However, expression analysis revealed that unlike in maize, the expression of *miR172* in *C. grandis* is not sex-biased (Figure 2-4F).

2.3.2 RNA sequencing, Trinity-based de novo transcriptome assembly and annotation using BLAST2GO and Trinotate

A total of 306575536 paired-end reads (150 bp) were obtained after sequencing all the twelve libraries on the Illumina NextSeq 500 platform. Subsequently, 186399131 good quality paired-end reads were used for *de novo* assembly of *Coccinia grandis* flower bud transcriptome (Table 2-1).

Table 2-1. RNA sequencing read counts and alignment statistics for all the samples used for *de novo* transcriptome assembly. Reproduced from Devani *et al.* (2017).

Sample name	Raw Reads	Cleaned reads	% Read pairs mapped concordantly
Male Early A	21719110	13216446	76.33%
Male Early B	22080977	12149633	74.84%
Female Early A	25607195	15634598	76.34%
Female Early B	25433955	15316996	75.94%
GyM-H Early A	27936206	17378706	76.67%
GyM-H Early B	27617808	17107330	76.69%
Ag-H Early A	26147527	15140736	71.02%
Ag-H Early B	25392540	15128539	69.35%
Male Middle A	25502209	15490027	76.70%
Male Middle B	25837400	15919727	75.27%
GyM-H Middle A	28465770	17441923	76.90%
GyM-H Middle B	24834839	16474470	77.43%
TOTAL	**306575536**	**186399131**	

Two different software were used for assembling the transcriptome namely Bridger and Trinity. The assembly generated by Bridger consisted of approximately ~153000 transcripts, whereas, the assembly generated by Trinity consisted of 467233 'Trinity Transcripts' clustering into 378860 'Trinity Genes' with an N50 value of 881 bp (Table 2-2, Additional file 1).

Table 2-2. Assembly statistics for *C. grandis* flower bud transcriptome generated using Trinity. Reproduced from Devani *et al.* (2017).

Parameter	Assembly statistics
Number of 'Trinity Transcripts'	467233
Number of 'Trinity Genes'	378860
Percent GC	38.96
Median contig length (bp)	347
Average contig length (bp)	606.45
N50 (bp)	881
Total assembled bases	283354298

For Trinity assembly, BUSCO output [C:89.8%(S:14.5%,D:75.3),F:5.0%,M:5.2%, n:1440] showed that out of 1440 BUSCOs for Plants dataset, 1293 full-length BUSCOs were detected in our *de novo*-assembled *Coccinia grandis* flower bud transcriptome indicating 89.8% completeness. Similarly, for Bridger assembly, BUSCO output was [C:91.7%[S:34.5%,D: 57.2%],F:3.3%,M:5.0%, n:1440] indicating 91.7 % completeness. Using TransRate, we were able to detect *C. grandis* homologs for 84 % (18039) of protein-coding primary transcripts of *C. sativus* in Trinity assembly, of which, 13430 reference sequences had at least 95 % of their bases covered by a CRB-BLAST hit (Table 2-3). Whereas, for Bridger assembly, similar analysis detected *C. grandis* homologs for 78 % (16507) of protein-coding primary transcripts of *C. sativus* in Trinity assembly, of which, 11955 reference sequences had at least 95 % of their bases covered by a CRB-BLAST hit (Table 2-3). We found that the assembly quality was comparable between the two software (Trinity and Bridger) when assessed by BUSCO and Transrate.

Table 2-3. *Coccinia grandis* flower bud transcriptome metrics calculated using TransRate. Protein-coding primary transcripts of *Cucumis sativus* were chosen as reference. Adapted from Devani *et al.* (2017).

Contig metrics	Trinity	Bridger	Comparative metrics	Trinity	Bridger
n seqs	467233	153044	CRBB hits	115567	54736
smallest	201	182	n contigs with CRBB	115567	54736
largest	17089	18881	p contigs with CRBB	0.25	0.36
n bases	2.83E+08	186170843	rbh per reference	5.37	2.55
mean len	606.45	1216.45	n refs with CRBB	18039	16507
n under 200	0	2	p refs with CRBB	0.84	0.77
n over 1k	68072	54433	cov25	17607	15880
n over 10k	44	211	p cov25	0.82	0.74
n with orf	70688	43151	cov50	16890	15105
mean orf percent	54.73	41.97	p cov50	0.79	0.7
n90	260	457	cov75	15744	13975
n70	441	1380	p cov75	0.73	0.65
n50	881	2441	cov85	14934	13262
n30	1713	3687	p cov85	0.69	0.62
n10	3162	6000	cov95	13430	11955
gc	0.39	0.39	p cov95	0.62	0.56
bases n	0	20	reference coverage	0.84	
proportion n	0	0			

The downstream analyses is much more streamlined for Trinity assembly as the software package comes bundled with the scripts required for transcript quantification, differential expression analysis and functional annotation, etc. Also, there is an active community for technical support in order to help with troubleshooting the problems that one might encounter while using Trinity. Hence, Trinity assembly was chosen for all the downstream analyses reported in this study. The transcripts of 200-399 bp size were found to be most abundant in the length distribution of transcripts assembled by Trinity (Figure 2-5). However, a higher proportion of transcripts with length around 1000-2000 bp had a BLAST hit compared to the proportion of smaller transcripts (Figure 2-5).
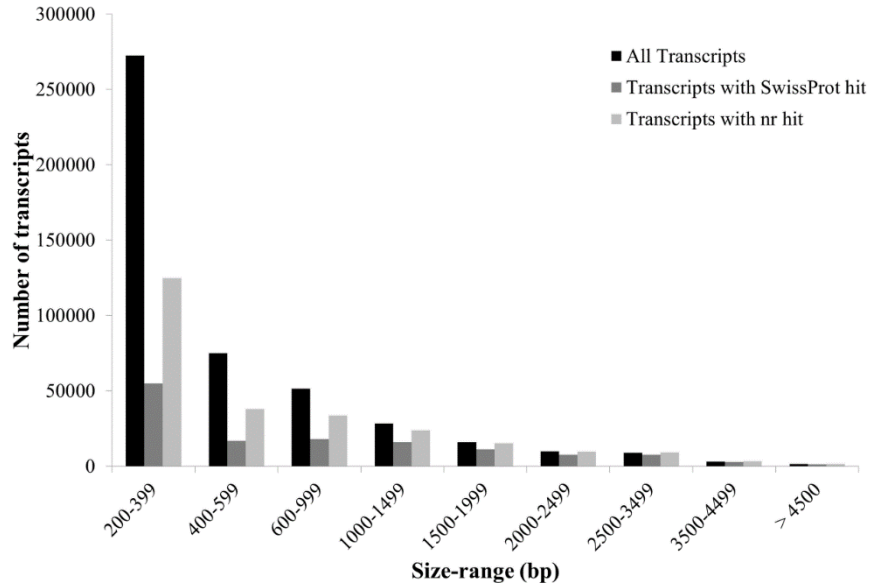
Figure 2-5. **Transcript size distribution for *C. grandis* flower bud transcriptome.** Reproduced from Devani *et al.* (2017).

Cleaned reads were mapped back to the transcriptome using bowtie2 with ~70 % or more reads from each library aligning concordantly (Table 2-1). An Ex90N50 statistic calculated using 80806 transcripts from the assembly (ignoring the rest of the transcripts with poor read coverage) was found to be 1784 bp (Figure 2-6). Altogether 8916 unique BLAST hits in the SwissProt database were represented by nearly full-length transcripts, having more than 70 % alignment coverage, and 12315 hits showed more than 50 % alignment coverage (Table 2-4).
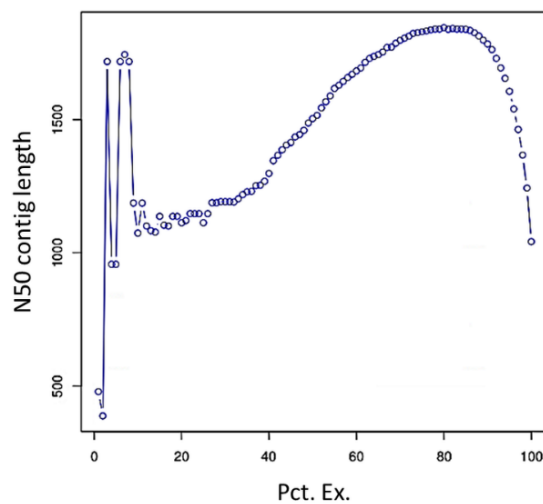
Figure 2-6. **ExN50 statistic for *C. grandis* flower *de novo* transcriptome assembly.** Reproduced from Devani *et al.* (2017).

Chapter 2: *De novo* transcriptomics approach for *C. grandis*

Table 2-4. Distribution of percent length coverage for the top matching SwissProt database entries. Reproduced from Devani *et al.* (2017).

#hit_pct_cov_bin	count_in_bin	>bin_below
100	5680	5680
90	1838	7518
80	1398	8916
70	1473	10389
60	1926	12315
50	2765	15080
40	5137	20217
30	10818	31035
20	28691	59726
10	46993	106719

Coccinia grandis flower bud transcripts were compared to plant protein sequences of the nr and SwissProt database resulting in 259200 and136663 transcripts having at least one hit from the respective database. Species distribution analysis of the BLAST hits showed that majority of these hits were from *Arabidopsis* and Rice for SwissProt database whereas for nr database most top hits were from cucumber and melon (Figure 2-7A, B). The number of transcripts annotated with various GO terms of biological process, molecular function, and cellular component categories are provided in Figure 2-7C.

Trinotate v3 pipeline was also used simultaneously for comprehensive functional annotation of the *Coccinia grandis* flower bud transcripts. Details regarding the SwissProt/TrEMBL BLAST hits, GO, KEGG and eggNOG mappings can be found in the Additional File 2. HMMER/PFAM predicted protein domains, as well as information regarding signal peptides and transmembrane domains could also be found in the Additional File 2. Taken together, we have assembled a good quality transcriptome for early and middle-staged flower buds of *Coccinia grandis* and comprehensively annotated the transcripts using well-established BLAST2GO and Trinotate pipelines.
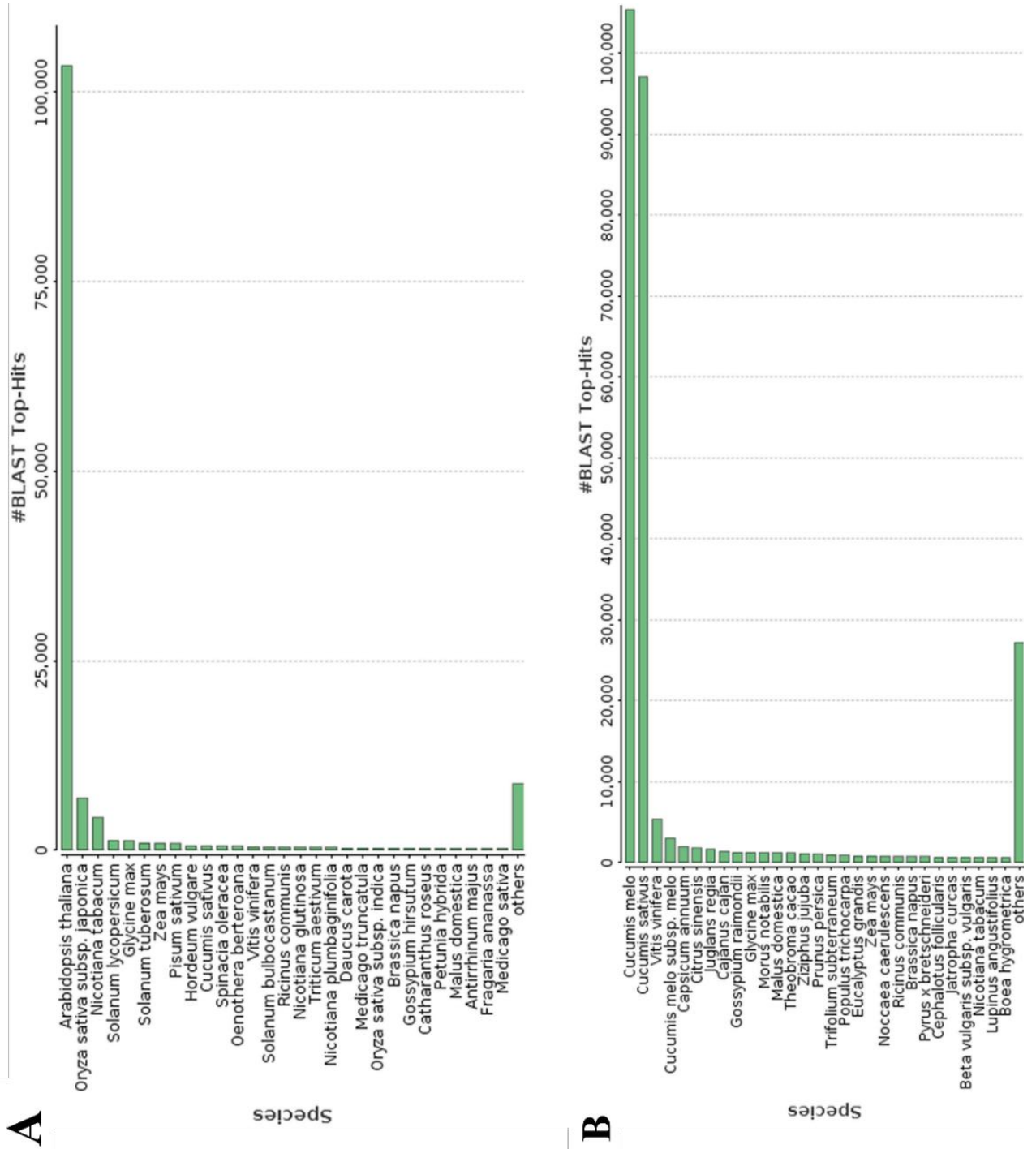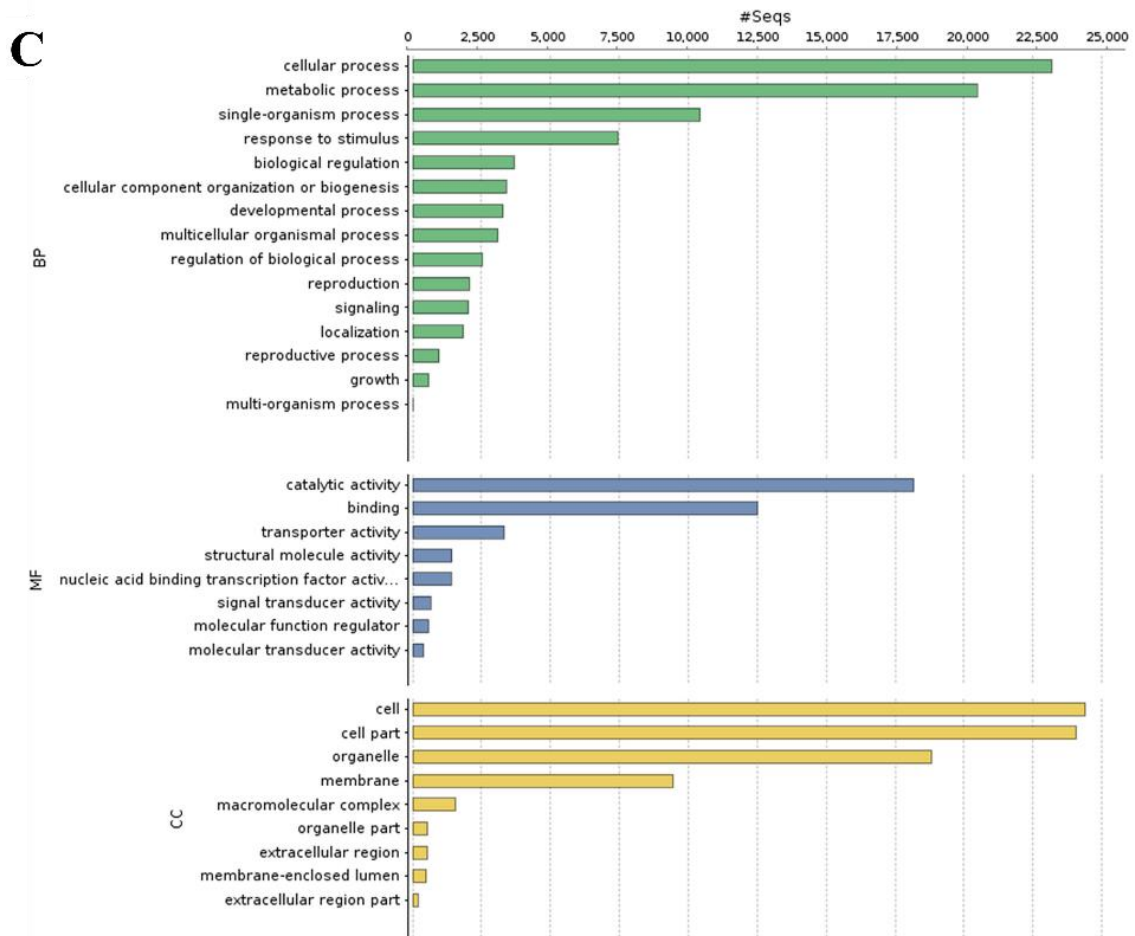
Figure 2-7. **BLAST2GO annotation of *C. grandis* flower bud transcriptome.** (A) BLAST Top-Hits species distribution when compared with Swiss-Prot database, (B) BLAST Top-Hits species distribution when compared with nr database, (C) GO category distribution of *C. grandis* flower bud transcriptome. Reproduced from Devani *et al.* (2017).

2.3.3 Differential Expression analysis reveals probable factors for pollen fertility and sex modification

Following transcriptome assembly and annotation, differential expression analysis was carried out. First of all, RSEM was used for transcript abundance estimation. Following which, we checked for the correlation between the replicates for all the samples using *PtR* script. PCA analysis and correlation matrix showed a good correlation between the replicate sets for each of the six samples (Figure 2-8). EdgeR was used to identify the differentially expressed transcripts for all the pairwise comparisons between the six samples (Table 2-5; Figure 2-9). Differentially expressed transcripts at a minimum fold change of 2^2 with p-values at most 1e-3 were extracted and GO enrichment analysis was performed (Additional File 3, Additional File 4).
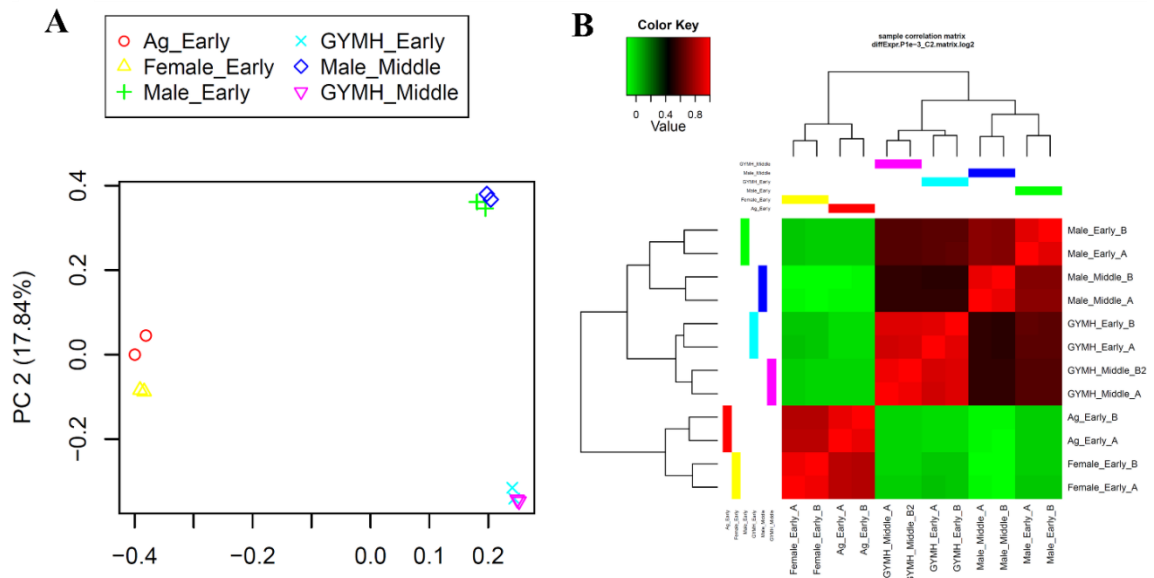
Figure 2-8. **Correlation analyses showing the relationship between samples and replicates.** (a) Principal component analysis and (b) correlation matrix showing relationship between all samples as well as replicates. Reproduced from Devani *et al.* (2017).
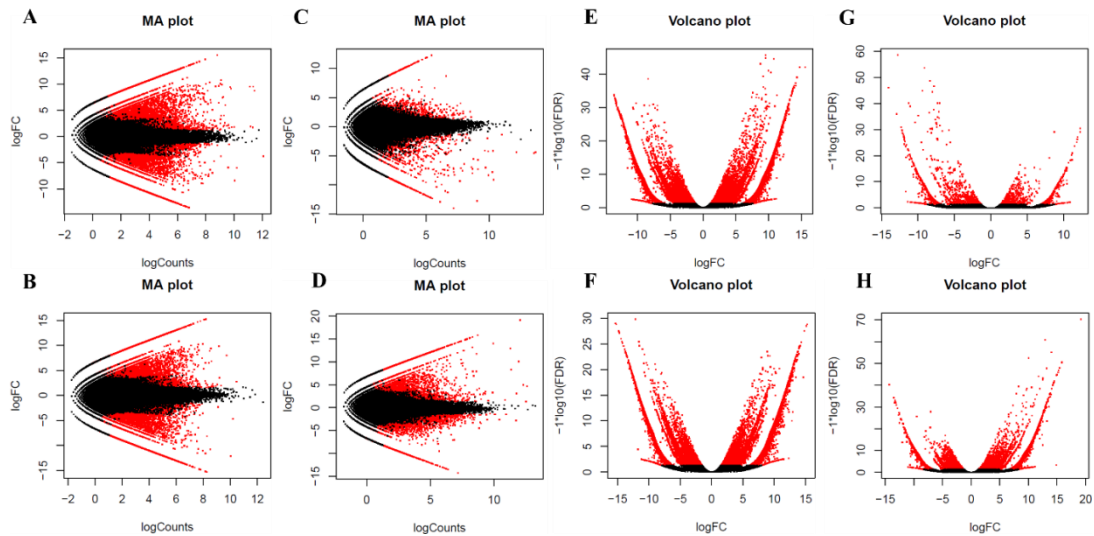
Figure 2-9. **Pairwise comparisons of transcript abundance.** MA plots showing average log fold change (logFC) vs average log of counts among (A) female (early-staged) vs. male (early-staged) transcripts, (B) female (early-staged) vs. GyM-H (early-staged) transcripts, (C) Ag-H (early-staged) vs. female (early-staged) transcripts and (D) GyM-H (middle-staged) vs. male (middle-staged) across replicates. Volcano plots showing differentially expressed transcripts in relation to False discovery rate (FDR) for (E) female (early-staged) vs. male (early-staged) transcripts, (F) female (early-staged) vs. GyM-H (early-staged) transcripts, (G) Ag-H (early-staged) vs. female (early-staged) transcripts and (H) GyM-H (middle-staged) vs. male (middle-staged). Features found DE at FDR < 0.05 are colored red. Features with P-values at most 1e-3 and at least 2^2 fold change are differentially expressed. Reproduced from Devani *et al.* (2017).

Chapter 2: *De novo* transcriptomics approach for *C. grandis*

Among all the comparisons, few interesting ones such as Ag_Early_vs_Female_Early, Female_Early_vs_Male_Early, and GYM-H_Middle_vs_Male_Middle had 3574, 35694 and 14954 differentially expressed transcripts respectively (Table 2-5, Figure 2-9).

Table 2-5. Number of differentially expressed transcripts for each pairwise comparison between the flower types. Transcripts that had p-values at most 1e-3 and were at least 2^2 fold were considered as differentially expressed. Reproduced from Devani *et al.* (2017).

Flower Buds Comparison	Number of DE transcripts
Ag_Early_vs_Female_Early	3574
Ag_Early_vs_GYMH_Early	34458
Ag_Early_vs_GYMH_Middle	38849
Ag_Early_vs_Male_Early	33863
Ag_Early_vs_Male_Middle	36923
Female_Early_vs_GYMH_Early	31886
Female_Early_vs_GYMH_Middle	39885
Female_Early_vs_Male_Early	35694
Female_Early_vs_Male_Middle	40477
GYMH_Early_vs_GYMH_Middle	816
GYMH_Early_vs_Male_Early	8659
GYMH_Early_vs_Male_Middle	11576
GYMH_Middle_vs_Male_Early	12357
GYMH_Middle_vs_Male_Middle	14954
Male_Early_vs_Male_Middle	4427

The DE features were partitioned into clusters with similar expression patterns (Figure 2-10; Figure 2-11). In the context of anther development, we identified several GO terms (GO:0080110, GO:0010208, GO:0010584, GO:0009555, GO:0055046, GO:0048658, GO:0048653) differentially enriched in male buds compared to female buds at an early stage of floral development.
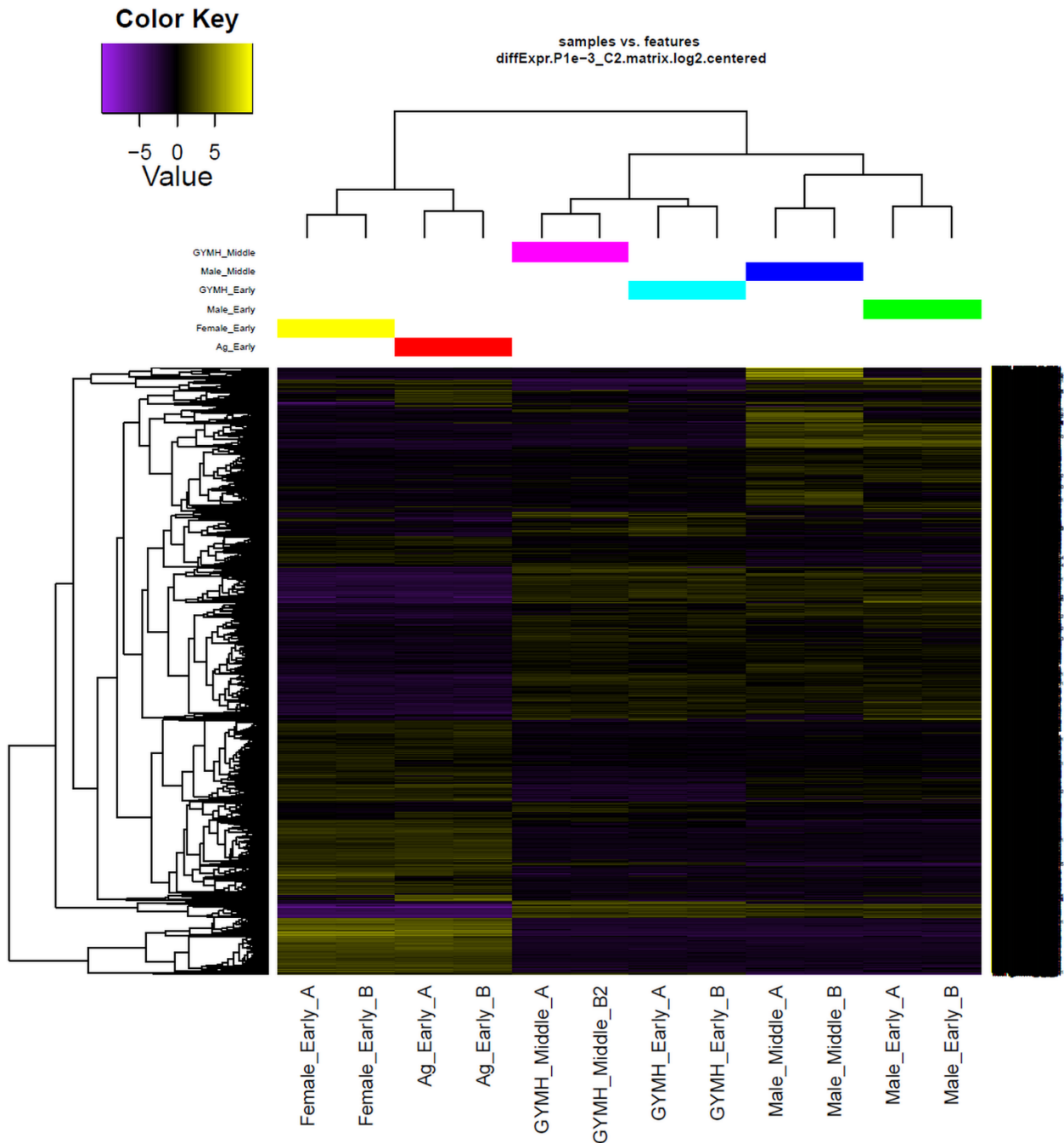
Figure 2-10. **Hierarchical clustering of differentially expressed transcripts and developmentally staged *C. grandis* flower bud samples.** Heatmap showing the relative expression levels of each transcript (rows) in each sample (columns). Rows and columns are hierarchically clustered. Expression values (FPKM) are log2 –transformed and then median-centred by transcript. Reproduced from Devani *et al.* (2017).

subcluster_1_log2_medianCentered_fpkm.matrix, 25612 tra   subcluster_2_log2_medianCentered_fpkm.matrix, 859 tra
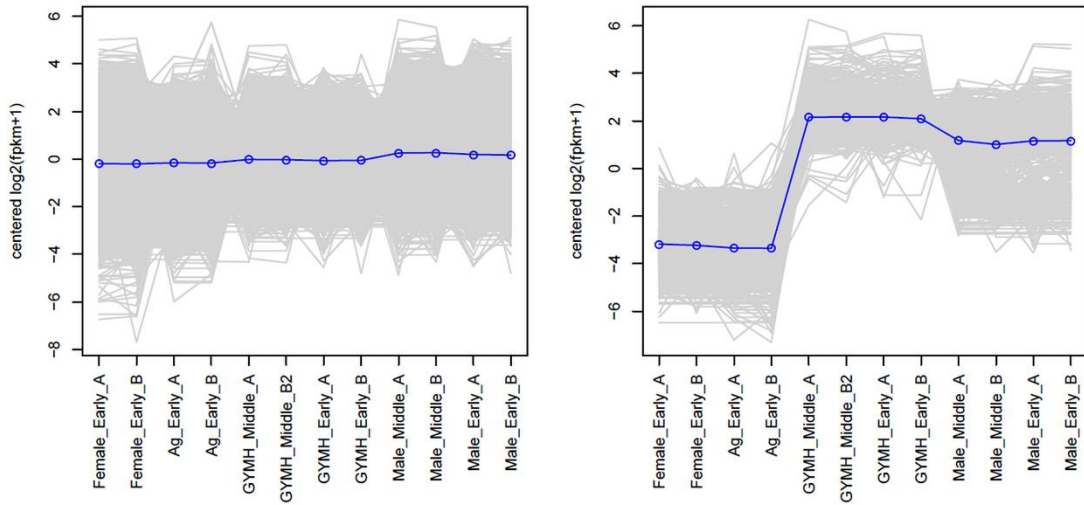
subcluster_3_log2_medianCentered_fpkm.matrix, 323 tra   subcluster_4_log2_medianCentered_fpkm.matrix, 2754 tra
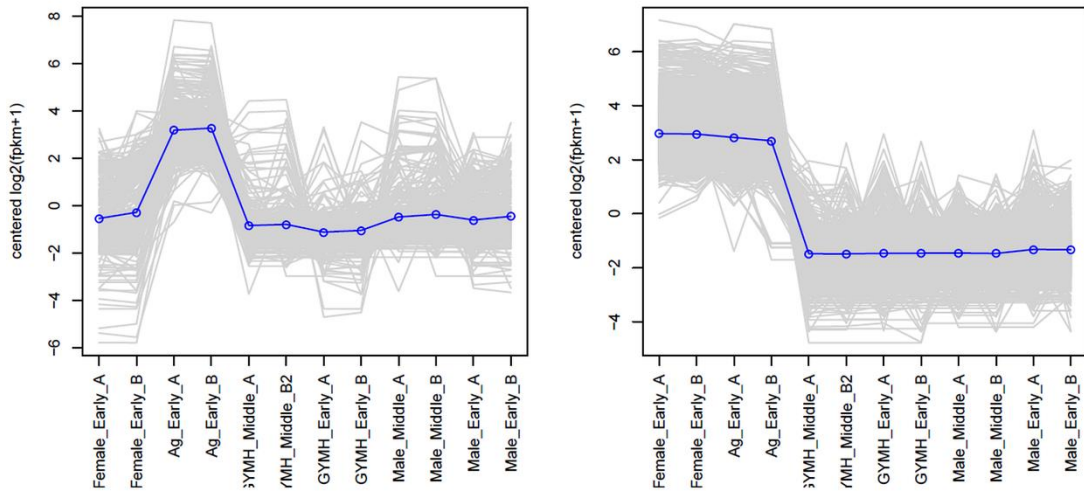
subcluster_5_log2_medianCentered_fpkm.matrix, 483 tra   subcluster_6_log2_medianCentered_fpkm.matrix, 120 tra

Chapter 2: *De novo* transcriptomics approach for *C. grandis*
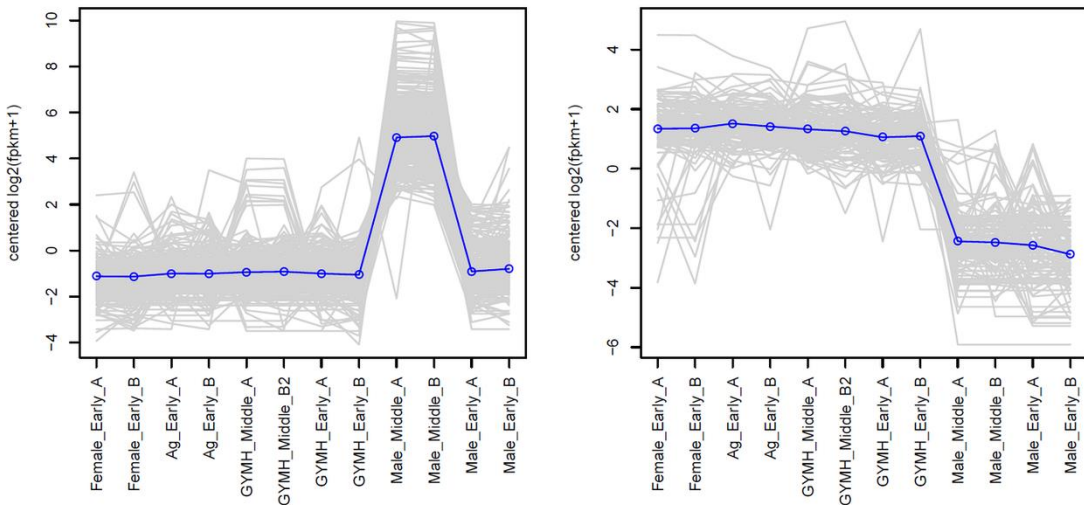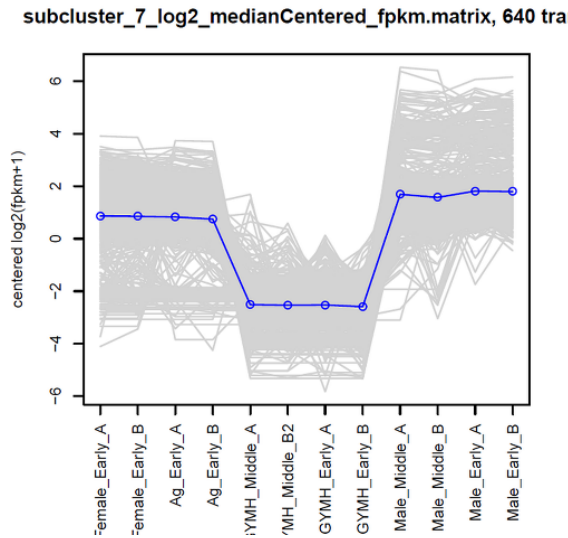
Figure 2-11. **Transcript clusters extracted from the hierarchical clustering with R.** X-axis: samples; y-axis: median-centred log$_2$(FPKM). Grey lines, individual transcripts; blue line, average expression values per cluster. Reproduced from Devani *et al.* (2017).

qRT-PCR was done to validate the results of differential expression analysis for a few interesting *Coccinia* homologs of *AMS* (ABORTED MICROSPORES), *CER3* (ECERIFERUM 3), *DEX1* (DEFECTIVE IN EXINE FORMATION 1), *DYT1* (DYSFUNCTIONAL TAPETUM 1), *EIL1* (ETHYLENE INSENSITIVE 3-like 1), *EMS1* (EXCESS MICROSPOROCYTES 1), *FER* (FERONIA), *MMD1* (MALE MEIOCYTE DEATH 1), *MS1* (MALE STERILITY 1), *SHT* (Spermidine hydroxycinnamoyl transferase), *TPD1* (TAPETUM DETERMINANT 1) and *ZAT3* (Zinc finger protein ZAT3). Expression profiles for these genes deduced by qRT-PCR revealed similar patterns to that seen from the digital DE analysis results (Figure 2-12).

Also, we have found GO terms related to pollen fertility enriched in the male buds (GO:0080092, GO:0009846, GO:0009860) compared to GyM-H and Ag-H buds, which had sterile pollens. Accordingly, expression profile for homologs of a number of genes involved in pollen tube development such as *CSLD1* (Cellulose synthase-like protein D1; TRINITY_DN92683_c0_g1_i1), *CDPKO* (Calcium-dependent protein kinase 24; TRINITY_DN93671_c0_g1_i3), *PME4* (Pectin methylesterase 4; TRINITY_DN14239_c0_g1_i1), *PME37* (Pectin methylesterase 37; TRINITY_DN3663_c0_g1_i1), *PPME1* (POLLEN SPECIFIC Pectin methylesterase 1;

Chapter 2: *De novo* transcriptomics approach for *C. grandis*

TRINITY_DN66415_c0_g1_i1) and *PTR52* (Protein NRT1/PTR FAMILY 2.8; TRINITY_DN112735_c0_g14_i3) were analysed and found to be enriched in middle-staged male buds similar to our digital expression profiles based on RNA-Seq data (Figure 2-13). Downregulation of ethylene signalling upon AgNO$_3$ treatment was evident as GO:0009723 (response to ethylene) and GO:0009873 (ethylene-activated signalling pathway) were depleted in AgNO$_3$ treated plant (Additional File 4). In order to validate this, we studied the expression profile of *Ethylene-responsive transcription factors*, *ERF5*, *ERF17,* and *EF102*. We found that all three *ERFs* were downregulated upon AgNO$_3$ treatment (Figure 2-14).
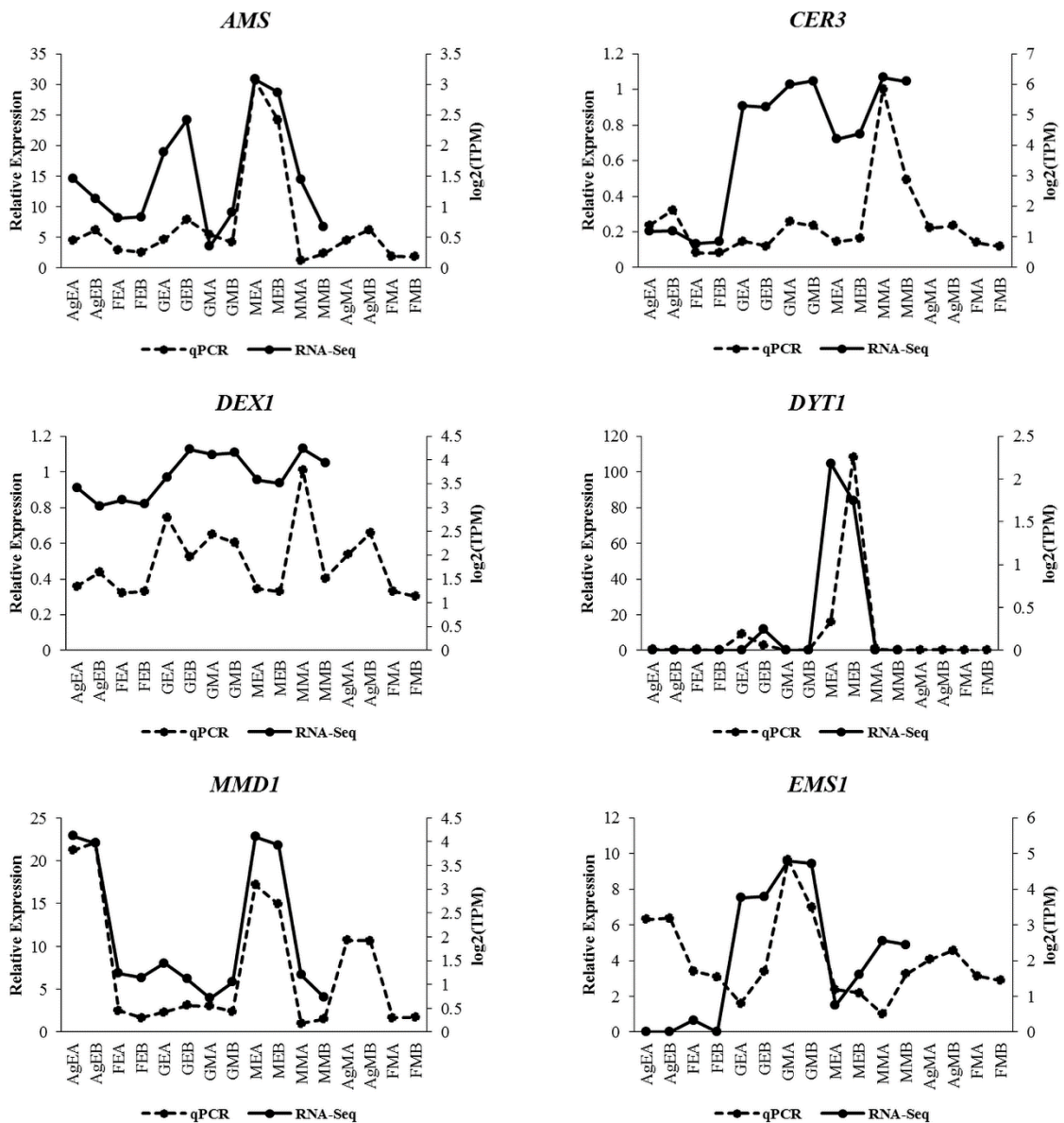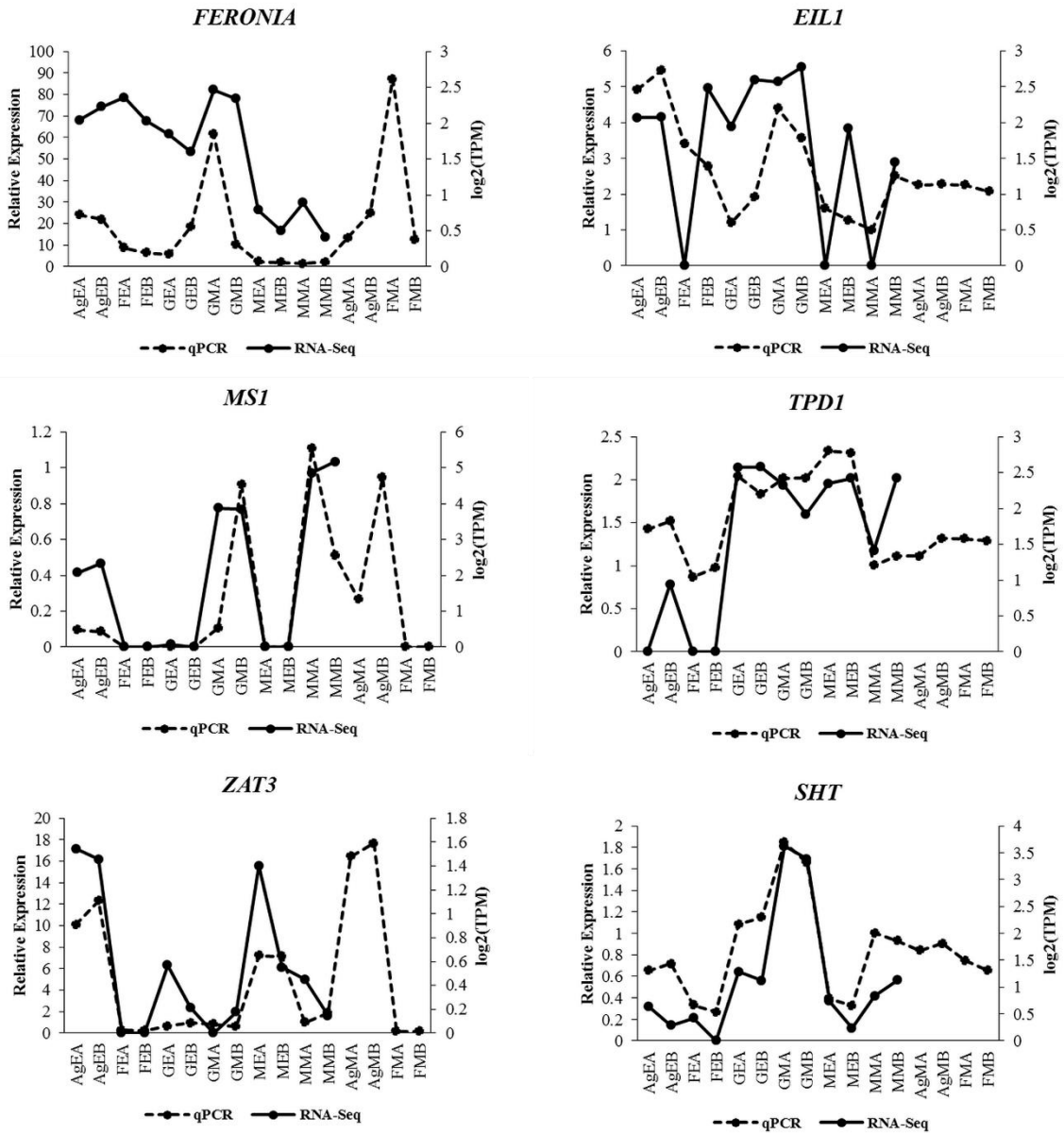
Figure 2-12. **Validation of selected DE genes by qRT-PCR with two biological replicates.** The relative expression in the sample of middle-staged male-A (MMA) was set to 1 for plotting the qRT-PCR data. AgEA, Early-staged Ag-H A; AgEB, Early-staged Ag-H B; FEA, Early-staged Female A; FEB, Early-staged Female B; GEA, Early-staged GyM-H A; GEB, Early-staged GyM-H B; GMA, Middle-staged GyM-H A; GMB, Middle-staged GyM-H B; MEA, Early-staged Male A; MEB, Early-staged Male B; MMA, Middle-staged Male A; MMB, Middle-staged Male B; AgMA, Middle-staged Ag-H A, AgMB, Middle-staged Ag-H B; FMA, Middle-staged Female A; FMB, Middle-staged Female B. Reproduced from Devani *et al.* (2017).
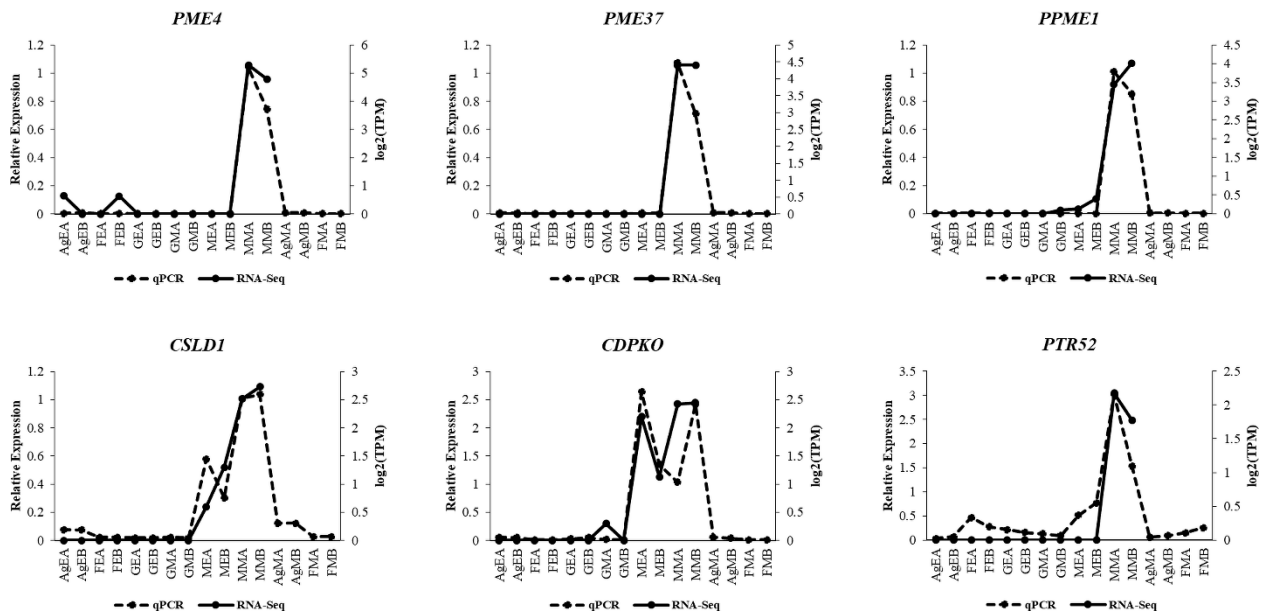
Figure 2-13. **qRT-PCR based expression analyses of selected genes involved in pollen tube development with two biological replicates.** The relative expression in the sample of middle-staged male-A (MMA) was set to 1 for plotting the qRT-PCR data. AgEA, Early-staged Ag-H A; AgEB, Early-staged Ag-H B; FEA, Early-staged Female A; FEB, Early-staged Female B; GEA, Early-staged GyM-H A; GEB, Early-staged GyM-H B; GMA, Middle-staged GyM-H A; GMB, Middle-staged GyM-H B; MEA, Early-staged Male A; MEB, Early-staged Male B; MMA, Middle-staged Male A; MMB, Middle-staged Male B; AgMA, Middle-staged Ag-H A, AgMB, Middle-staged Ag-H B; FMA, Middle-staged Female A; FMB, Middle-staged Female B. Reproduced from Devani *et al.* (2017).

Figure 2-14. **qRT-PCR based expression analyses of selected *Ethylene-responsive transcription factors* (*ERFs*) with two biological replicates.** The relative expression in the sample of middle-staged male-A (MMA) was set to 1 for plotting the qRT-PCR data. AgEA, Early-staged Ag-H A; AgEB, Early-staged Ag-H B; FEA, Early-staged Female A; FEB, Early-staged Female B; GEA, Early-staged GyM-H A; GEB, Early-staged GyM-H B; GMA, Middle-staged GyM-H A; GMB, Middle-staged GyM-H B; MEA, Early-staged Male A; MEB, Early-staged Male B; MMA, Middle-staged Male A; MMB, Middle-staged Male B; AgMA, Middle-staged Ag-H A, AgMB, Middle-staged Ag-H B; FMA, Middle-staged Female A; FMB, Middle-staged Female B. Reproduced from Devani *et al.* (2017).

Chapter 2: *De novo* transcriptomics approach for *C. grandis*

2.4 Discussion

Genetic basis of sex determination and differentiation is not well studied in *C. grandis*. Identification and investigation of sex-linked genes would lead to better understanding of dioecy in plants, and this can be achieved by whole genome sequencing approach. However, sex-determining genes are most likely linked to non-recombining regions of Y-chromosome, which are difficult to assemble from sequence data (Muyle *et al.*, 2016). An alternative approach is to use comparative transcriptomics to identify sex-biased genes that could play a role in sex differentiation and determination (Muyle *et al.*, 2012). Further, the presence of mutations and SNPs in sex-biased genes can provide insights regarding the evolution of dioecy. Using this approach, we have assembled and annotated a *de novo* transcriptome from the flower buds of dioecious, gynomonoecious and AgNO$_3$ treated female *C. grandis.* We have identified differentially expressed genes which might be playing a role in stamen arrest of female flowers. Also, we have analysed the genes that were differentially expressed upon AgNO$_3$ treatment on female plants promoting stamen development. Finally, we have compared middle-staged male (bearing fertile pollens) and GyM-H buds (bearing sterile pollens) to study the genes involved in pollen maturation and fertility.

2.4.1 Differential expression of stamen developmental genes and arrest of stamen growth in female flowers

At the early stages (stages 3-4) of flower development in female *C. grandis,* both carpel and stamen organs are initiated simultaneously. However, stamen growth gets arrested during the course of development (stages 4-5) resulting in a female flower with rudimentary stamens. In contrast, no sign of carpel primordia was observed during the histological study of flower development in male *C. grandis* as described in our previous report (Ghadge *et al.*, 2014). The molecular players involved in stamen initiation and development process are well characterized in the hermaphrodite plant *Arabidopsis*. In order to identify the stage at which stamen growth gets arrested, *Coccinia grandis* homologs of *Arabidopsis* stamen development genes were identified from the *de novo*-assembled transcriptome. Among genes involved in stamen initiation, *Pistillata* (*CgPI*, TRINITY_DN71631_c0_g1_i1) was found to be expressed in a male-biased fashion (Additional File 3). Expression profiling of *CgPI* by qRT-PCR also showed

male-biased expression profile (Figure 2-4A). *Pistillata* has been shown to specify stamen identity in *Arabidopsis* (Krizek and Meyerowitz, 1996) (Table 2-6).

Table 2-6. Digital Expression profile for genes in anther developmental pathway. Reproduced from Devani *et al.* (2017).

Stage of stamen development	Genes	Expression pattern
Stamen Primordia Initiation	AG	Unbiased
	CLV1/CLV2	Unbiased/male-biased
	PI	Male-biased
	AP3	Unclear homolog
	JAG	Male-biased
Archesporial initiation	BAM1/BAM2	Unclear homolog
	SPL/NZZ	Unclear homolog
Tapetal Development	EMS1	Male-biased
	SERK1/2	Male-specific
	TPD1	Male-biased
	RPK2	Male-specific
	TDF	Male-biased
	DYT1	Unbiased
	bHLH89	Male-biased
Mature Pollen Formation	AMS	Male-biased
	MS1	Male-biased
	MS2	Male-biased
	MIA	Unbiased
	LAP3	Unbiased
	LAP5	Male-biased

Further, EXCESS MICROSPOROCYTES 1 (EMS1) has been shown to interact with TAPETUM DETERMINANT 1 (TPD1) regulating specification of reproductive as well as somatic cells in *Arabidopsis* anthers (Jia *et al.*, 2008). Differential expression analyses revealed

that homologs of both *EMS1* (TRINITY_DN106236_c0_g4_i1) and *TPD1* (TRINITY_DN116795_c2_g1_i3) were enriched in male flowers compared to female flowers (Table 2-6; Additional File 3, Figure 2-12). *DYSFUNCTIONAL TAPETUM 1 (DYT1)* plays an important role in tapetum development by regulating the expression of *DEFECTIVE IN TAPETAL DEVELOPMENT AND FUNCTION 1* (*TDF1*) in *Arabidopsis* (Gu *et al.*, 2014). Also, DYT1 is known to interact with Basic helix-loop-helix protein 89 (bHLH89) which is highly expressed in anthers and required for normal anther development and male fertility (Zhu *et al.*, 2015). *TDF1* homolog (TRINITY_DN97604_c1_g7_i1) as well as *bHLH89* homolog (TRINITY_DN85771_c0_g1_i1) showed male-biased expression in *C. grandis* (Table 2-6; Additional File 3). Differential regulation of these genes related to stamen development explains the possible cause for early stamen arrest in female flowers of *C. grandis*.

According to recent reports from monoecious cucurbits like melon, cucumber, and watermelon, ethylene plays a major role in sex determination by inhibiting stamen development process (Boualem *et al.*, 2008, Boualem *et al.*, 2009, Boualem *et al.*, 2015, Boualem *et al.*, 2016). We found that compared to male, GO:0009723 (response to ethylene) was enriched in female buds indicating a potential role of ethylene in sex determination of *C. grandis* (Additional File 4).

2.4.2 *AgNO₃ treatment on female plant releases the stamen inhibition*

Female plants of *C. grandis* bear flowers with fused carpels and rudimentary stamens. Earlier, we have shown that foliar spray of 35 mM AgNO₃ on the female plant of *C. grandis* promotes further development of the rudimentary stamens (Ghadge *et al.*, 2014). In the current study, gene expression profiles for early-staged Ag-H flower buds were compared with female buds (Table 2-5; Figure 2-9C, G). Ag$^+$ ions are known to inhibit responses to ethylene, a gaseous plant hormone (Beyer, 1976). Also, silver compounds have been shown to induce maleness by promoting stamen development in many monoecious and dioecious species (Sarath and Mohan Ram, 1979, Yin and Quinn, 1995, Law *et al.*, 2002). No other inhibitors of ethylene biosynthesis or signalling could induce the stamen development in *Silene latifolia,* suggesting that ethylene signalling might not be the only pathway that gets affected upon application of silver thiosulphate (Law *et al.*, 2002). In contrast to *Silene latifolia*, AVG (aminoethoxyvinylglycine), an inhibitor of ethylene-biosynthesis has been shown to induce male

flowers in gynoecious muskmelon similar to silver compounds (Owens *et al.*, 1980). Considering the role of 1-aminocyclopropane-1-carboxylate synthase (ACS, an enzyme involved in ethylene biosynthesis) in sex determination of many other members of Cucurbitaceae, an ethylene-mediated effect of $AgNO_3$ seems more likely to be involved in the modification of sex in *C. grandis* (Zhang *et al.*, 2014b, Zhang *et al.*, 2014c). In accordance with this, we observed that the expression of *CgACS* is female-biased in *C. grandis,* similar to monoecious cucurbits (Figure 2-2).

In our study, GO:0009723 (response to ethylene) and GO:0009873 (ethylene-activated signaling pathway) were enriched in female buds compared to Ag-H buds (Additional File 4). Transcripts for genes such as *Ethylene-responsive transcription factors, ERF5* (TRINITY_DN102355_c3_g13_i1), *ERF17* (TRINITY_DN80749_c0_g6_i1), EF109 (TRINITY_DN87049_c0_g1_i1), EF102 (TRINITY_DN90257_c1_g2_i1), ERF99 (TRINITY_DN93821_c0_g1_i2), ERF60 (TRINITY_DN93262_c1_g6_i2) and *ERF78 (TRINITY_DN98503_c3_g1_i1)* were downregulated in Ag-H buds indicating impaired ethylene signalling (Additional File 3). Additionally, qRT-PCR based expression pattern analysis for *ERF5*, *ERF17* and *EF102* genes clearly showed the suppression of ethylene responses by $AgNO_3$ (Figure 2-14).

Downregulation of ethylene signalling in Ag-H buds was correlated with the promotion of stamen growth. GO:0048655 (anther wall tapetum morphogenesis), GO:0048657 (anther wall tapetum cell differentiation), GO:0048658 (anther wall tapetum development) were seen to be enriched in early-staged Ag-H buds compared to female buds (Additional File 4). *C. grandis* homologs of *MS1*, *MMD1* (TRINITY_DN109512_c4_g3_i1, TRINITY_DN108927_c0_g6_i1), *ZAT3* (TRINITY_DN108658_c0_g2_i1) and *AMS* (TRINITY_DN116105_c0_g2_i1) genes which play important roles in tapetum and pollen development of *Arabidopsis* flowers were upregulated upon $AgNO_3$ treatment indicating promotion of stamen growth (Ito and Shinozaki, 2002, Sorensen *et al.*, 2003, Yang *et al.*, 2003, Ito *et al.*, 2007, Yang *et al.*, 2007, Borg *et al.*, 2011) (Additional File 3; Figure 2-12). *MYB35* (TRINITY_DN92649_c0_g7_i1), which was proposed as a putative sex-determining gene in *Asparagus* was also found to be upregulated in Ag-H buds (Tsugama *et al.*, 2017) (Additional File 3). Apart from that, gene ontology terms related to pollen wall assembly (GO:0010208), pollen exine formation (GO:0010584),

sporopollenin biosynthetic process (GO:0080110), pollen development (GO:0009555) and pollen sperm cell differentiation (GO:0048235) were also enriched in Ag-H buds (Additional File 4). Further, we noticed that *Ethylene-responsive transcription factors* (ERFs) were not affected in GyM-H buds as compared to female buds suggesting that stamen development in GyM-H flower buds might be regulated by some other mechanism evading ethylene signalling inhibition.

2.4.3 Transcripts governing pollen fertility are depleted in GyM-H and Ag-H flower buds

C. grandis is one of the few species in which the presence of heteromorphic sex chromosomes is reported. The large Y-chromosome present in males might play a major role in sex determination. The GyM form of *C. grandis* included in the current study does not have Y-chromosome (Ghadge *et al.*, 2014). GyM-H flowers still develop full-sized stamens despite lacking Y-chromosome. Similarly, AgNO$_3$ treatment induces stamen development in female plants having XX sex chromosomes. However, the pollens from GyM-H and Ag-H flowers buds were found to be sterile unlike the pollens from male buds (Ghadge *et al.*, 2014). Differential expression analysis revealed that gene ontology terms for pollen tube (GO:0090406), pollen germination (GO:0009846), regulation of pollen tube growth (GO:0080092), pollen tube growth (GO:0009860) and microsporogenesis (GO:0009556) were enriched in middle-staged male buds compared to middle-staged GyM-H buds (Additional File 4).

GAUTE plays an important role in pollen tube wall biosynthesis in *Arabidopsis* (Wang *et al.*, 2013). TRINITY_DN111340_c1_g1_i6, which showed similarity with *GAUTE* was enriched in male buds compared to GyM-H buds. Unlike most other plant cell walls, pollen tube wall does not contain callose or cellulose. Pectin methylesterases (PMEs) have been shown to play a very important role in the growth of pollen tubes (Jiang *et al.*, 2005, Tian *et al.*, 2006, Leroux *et al.*, 2015). *PME4* (TRINITY_DN14239_c0_g1_i1), *PME37* (TRINITY_DN3663_c0_g1_i1) and *PPME1* (TRINITY_DN66415_c0_g1_i1, TRINITY_DN71598_c0_g2_i1) were downregulated in GyM-H buds compared to male buds (Additional File 3). This could be a possible cause for pollens from GyM-H not forming pollen tubes. Zhan *et al.* (2015) has shown that *IPMKB* (Inositol polyphosphate multikinase beta) is an important factor for pollen development. We have found that TRINITY_DN96290_c0_g3_i2 transcript matching to Arabidopsis *IPMKB* (*AtIpk2beta)* was downregulated in GyM-H compared to male buds. Earlier, several reports have

demonstrated that *MALE STERILITY 1* (*MS1*) gene of *Arabidopsis* expresses in tapetal cells and plays an important role in pollen maturation (Ito and Shinozaki, 2002, Ito *et al.*, 2007, Yang *et al.*, 2007). *C. grandis* homolog of *MS1*, TRINITY_DN109512_c4_g3_i1 was expressed in a male-biased manner (Additional File 3; Figure 2-12). Similarly, homologs of genes important for pollen tube growth such as *CSLD1* (TRINITY_DN92683_c0_g1_i1), *CDPKO* (TRINITY_DN93671_c0_g1_i3), *NRX1* (TRINITY_DN106708_c1_g2_i3), *PTR52* (TRINITY_DN112735_c0_g14_i3; TRINITY_DN112735_c0_g3_i1), *TAF6* (TRINITY_DN96231_c1_g1_i2) and *CALS5* (TRINITY_DN113564_c1_g1_i1) were enriched in male (Dong *et al.*, 2005, Lago *et al.*, 2005, Bernal *et al.*, 2008, Qin *et al.*, 2009, Xie *et al.*, 2010, Zhao *et al.*, 2013) (Figure 2-13, Additional File 3). Genes involved in pollen exine formation such as *FACR2*/*MS2* (TRINITY_DN74585_c1_g5_i3), *EA6* (TRINITY_DN76274_c1_g1_i1), *C70A2/DEX2* (TRINITY_DN99059_c0_g1_i1) were also upregulated in male (Morant *et al.*, 2007, Dobritsa *et al.*, 2011, Xiong *et al.*, 2016). *EMS1* (TRINITY_DN89942_c0_g7_i1), *SERK1* (TRINITY_DN108624_c1_g7_i5), *JASON* (TRINITY_DN83440_c0_g1_i1), *RPK2* (TRINITY_DN113423_c0_g1_i4), which are essential for microsporogenesis and pollen maturation were observed to be expressed at significantly higher levels in middle-staged male buds compared to GyM-H buds. (Zhao *et al.*, 2002, Albrecht *et al.*, 2005, Mizuno *et al.*, 2007, De Storme and Geelen, 2011) (Additional File 3; Figure 2-12)

Expression profiling for homologs of *MS1*, *EMS1*, *DYT1*, *PME4*, *PME37*, *PPME1*, *CSLD1*, *CDPKO*, *and PTR52* was studied by qRT-PCR for all the tissue samples including middle-staged Ag-H buds (Figure 2-12; Figure 2-13; Table 2-6). Transcripts for all these homologs were downregulated in Ag-H buds and GyM-H buds, suggesting a male-biased expression pattern implicating the reason for pollen sterility in Ag-H and GyM-H buds.

2.5 Conclusions

De novo-assembled transcriptome developed from RNA-Seq of different sexual phenotypes has enabled identification of *C. grandis* homologs of many genes known to be involved in flower development in species such as *Arabidopsis*, melon, and cucumber. We found out that many genes involved in stamen initiation, tapetal development, and pollen maturation were downregulated in female buds compared to male buds. Interestingly, ethylene response-related genes were upregulated in female buds compared to male buds indicating a probable role

of ethylene in stamen suppression similar to monoecious cucurbits such as melon and cucumber. We speculate that the Y-chromosome might express genes that inhibit ethylene signalling or suppress the carpel development, the site of ethylene production leading to the formation of stamens in male flowers. This was supported by the observation that AgNO₃ treatment suppressed ethylene responses and induced stamen development in female flowers of *C. grandis*. However, the pollens produced by Ag-H flowers were sterile indicating a decisive role of Y-chromosome in determining maleness. In accordance with this, the transcripts involved in pollen maturation, pollen germination, and pollen tube elongation were depleted in middle-staged GyM-H buds compared to male buds. This could be because of the absence of Y-chromosome in GyM plant. Altogether, differentially expressed genes identified in this study could shed light on the probable mechanisms of sex determination, differentiation, and modification in *C. grandis*.

2.6 Availability of data and material

The raw sequencing data were deposited in the NCBI Short Read Archive (SRA) database (http://www.ncbi.nlm.nih.gov/sra/) under the accession number SRP111347.

2.7 Supporting information

Additional File 1: *De novo*-assembled *Coccinia grandis* floral bud transcriptome. Available at

https://figshare.com/s/8bca31a89a1ff6000c09

Additional File 2: Annotation report for the *de novo*-assembled *C. grandis* flower bud transcriptome generated using Trinotate. Available at https://figshare.com/s/3068205f6e5b8095ec76

Additional File 3: Expression matrix of transcripts with P-values at most 1e-3 and at least 2^2 fold differential expression. Available at https://figshare.com/s/d4d910b3f0b0e49bb8f7

Additional File 4: Gene ontology (GO) enrichment analysis for differentially expressed transcripts in each pairwise comparison. Available at https://figshare.com/s/e51699d3b9f4bbe5f17d

Results presented in this chapter have been published in the following research article:

Devani R.S., Sinha S., Banerjee J., Sinha R.K., Bendahmane A. and Banerjee A.K. (2017). *De novo* transcriptome assembly from flower buds of dioecious, gynomonoecious and chemically masculinized female *Coccinia grandis* reveals genes associated with sex expression and modification. **BMC Plant Biology.** DOI:10.1186/s12870-017-1187-z

Chapter 3: Flower bud proteome reveals modulation of sex-biased proteins potentially associated with sex expression and modification in dioecious *Coccinia grandis*

3.1 Background

Flowering plants show three major sexual systems *viz.* hermaphroditism, monoecy and dioecy. Around 90 % of the angiosperm species are hermaphroditic bearing perfect flowers having both male as well as female reproductive sex organs (Yampolsky, 1922). Monoecy exists at a frequency of ~5-6 % in angiosperms, wherein unisexual male and female flowers are produced on the same individual plant. Remaining ~ 5-6 % angiosperm species are dioecious, having separate unisexual plants bearing either only male or only female flowers (Renner and Ricklefs, 1995, Charlesworth, 2002, Renner, 2014). Dioecious species show patchy phylogenetic distribution and are reported in around three-fourths of the angiosperm families. This indicates that the evolution of dioecy has occurred multiple times in different families independently and hence, the molecular mechanisms of sex determination might vary between distant dioecious species and is a matter of great research interest (Ainsworth *et al.*, 1997b, Guttman and Charlesworth, 1998, Ainsworth, 2000). Out of ~15,600 known dioecious species in 175 families, a very limited number have evolved sex chromosomes. Plant sex chromosomes have been reported in around just 40 species till date (Ming *et al.*, 2011). Sex determination in rest of the dioecious species may be governed by a single locus or multiple loci on autosomes, which may or may not be linked (Zhang *et al.*, 2014c, Yang *et al.*, 2015). The mechanism of sex determination in plants can be complex, and it may also involve environmental factors apart from genetic factors (Dellaporta and Calderon-Urrea, 1993). *Coccinia grandis* (L.) Voigt, is a dioecious member of Cucurbitaceae, a family known for its diverse sexual systems (Kouonon *et al.*, 2009). Detailed description of *C. grandis* sexual forms and $AgNO_3$-mediated sex modification can be referred in section 2.1 in this thesis.

At present, our knowledge about the sex determination mechanisms in plants is fairly limited. Major limitation for studying mechanisms of sex determination in plants is that majority of the dioecious plants are non-model organisms without genome sequence. Hence, the rate at which sex-linked genes are identified from dioecious plant species is very low (Muyle *et al.*, 2012). However, the advent of NGS (next-generation sequencing) technologies has enabled the high-throughput identification of sex-biased genes from dioecious plant species in recent times (Muyle *et al.*, 2012). Also, advanced proteomic approaches may lead to the identification of

novel sex-linked proteins that can eventually expand our understanding in evolutionary, developmental and molecular mechanism associated with sex determination and modification.

Muyle *et al.* (2012) took an RNA-Seq approach to identify sex-biased gene expression in *Silene latifolia* and demonstrated the dosage compensation in plants for the first time. Similarly, other transcriptome studies in *Silene* have helped in understanding the Y chromosome degeneration and identification of new sex-linked genes (Bergero and Charlesworth, 2011, Chibalina and Filatov, 2011).Another study in persimmon (*Diospyros lotus*) showed that *OGI* (*Oppressor of MeGI*), a Y-chromosome-encoded small RNA governs pollen fertility by targeting a homeodomain transcription factor *MeGI* in a dose-dependent manner (Akagi *et al.*, 2014). Comparative *de novo* transcriptomics approach taken in *Asparagus* resulted in identification of genes involved in pollen microspore and tapetum development that were expressed specifically in male flowers (Harkess *et al.*, 2015). Similar transcriptomic studies have also been carried out in papaya and cucumber that has led to better understanding of sex determination (Guo *et al.*, 2010, Wu *et al.*, 2010, Urasaki *et al.*, 2012). It can be noted that recently, we have carried out a *de novo* transcriptome profiling from male, female, GyM-H and Ag-H flower buds of *C. grandis* and identified many sex-biased genes that can provide crucial insights into stamen arrest, pollen fertility and sex modification mechanism (Devani *et al.*, 2017).

By using transcriptomics over genomics, it is possible to capture differential gene regulations that may arise due to changes in environmental cues/signals or as in case of dioecious plants, the differential gene regulations based on the sex of the plant can be identified. However, transcriptomics has a disadvantage that the mRNA abundance may not always accurately reflect protein levels, which are the end-products in the realization of hereditary information carrying out various structural and functional duties (Gygi *et al.*, 1999, Zhang *et al.*, 2014a). Moreover, only proteomics can capture the post-translational modifications which are very well-known to control protein functions. However, proteomics study generally depends on the availability of genomic/transcriptomic data which is limited for most of the dioecious species. Hence, there are only few studies which have employed proteomics approach in order to understand sex determination and differentiation in dioecious plants (Yang *et al.*, 2015). One such study carried out in *Asparagus* identified differentially accumulated proteins in the form of spots on 2-D gels from flowers of male and female plants (Bracale *et al.*, 1990). Another study in *Pistacia vera*

purified a 27 kDa glycoprotein specific to female inflorescence (Golan-Goldhirsh *et al.*, 1998). Proteomic approaches has also been attempted to identify sex-linked proteins in *Ginkgo biloba*. A 28 kDa protein specific to male and a 36 kDa as well as 92 kDa proteins specific to female inflorescences has been identified (Yang *et al.*, 2015). Recently, Manzano *et al.* (2017) showed that overexpression of aerolysin-like protein from the dioecious plant *Rumex acetosa* induces male sterility in transgenic tobacco.

In order to decipher differences in protein abundances among flower buds of male, female, gynomonoecious and chemically masculinized female plants of *C. grandis*, total proteome profiling was carried out at early and middle stages of flower development. Proteins were identified using IDA-derived data using *in silico* translated in-house *C. grandis* flower bud transcriptome (Devani *et al.*, 2017) database followed by label-free quantification using SWATH-MS data. Differential protein abundance among various sex forms of *C. grandis* was studied to identify players potentially involved in sex expression and modification. Our study has identified key proteins potentially involved in the processes of stamen arrest, AgNO$_3$-mediated sex modification and pollen fertility in dioecious *C. grandis*.

3.2 Materials and methods

3.2.1 *Flower bud collection*

Both early and middle stages flower buds of male, female, gynomonoecious and chemically masculinized female plants of *C. grandis* were harvested as per our previous protocol (Devani *et al.*, 2017). Overall, experimental plan is depicted schematically in Figure 3-1. In early-staged male flower buds, only stamens are present with no sign of carpel initials, while same staged female flowers (stages 3-4) have both carpel and stamen primordia (Chapter 2, Figure 2-1). Stamen growth in female flowers gets arrested around stages 4–5 (Chapter 2, Figure 2-1). In the hermaphrodite flowers of gynomonoecious plant, however, both stamens and carpels develop simultaneously during early as well as middle stages of development. The selection of early-staged flower buds was carried out such that the event of stamen inhibition in female flowers can be analysed. Similarly, middle-staged flower buds were chosen such that meiosis-stage and pollen maturation event can be investigated.
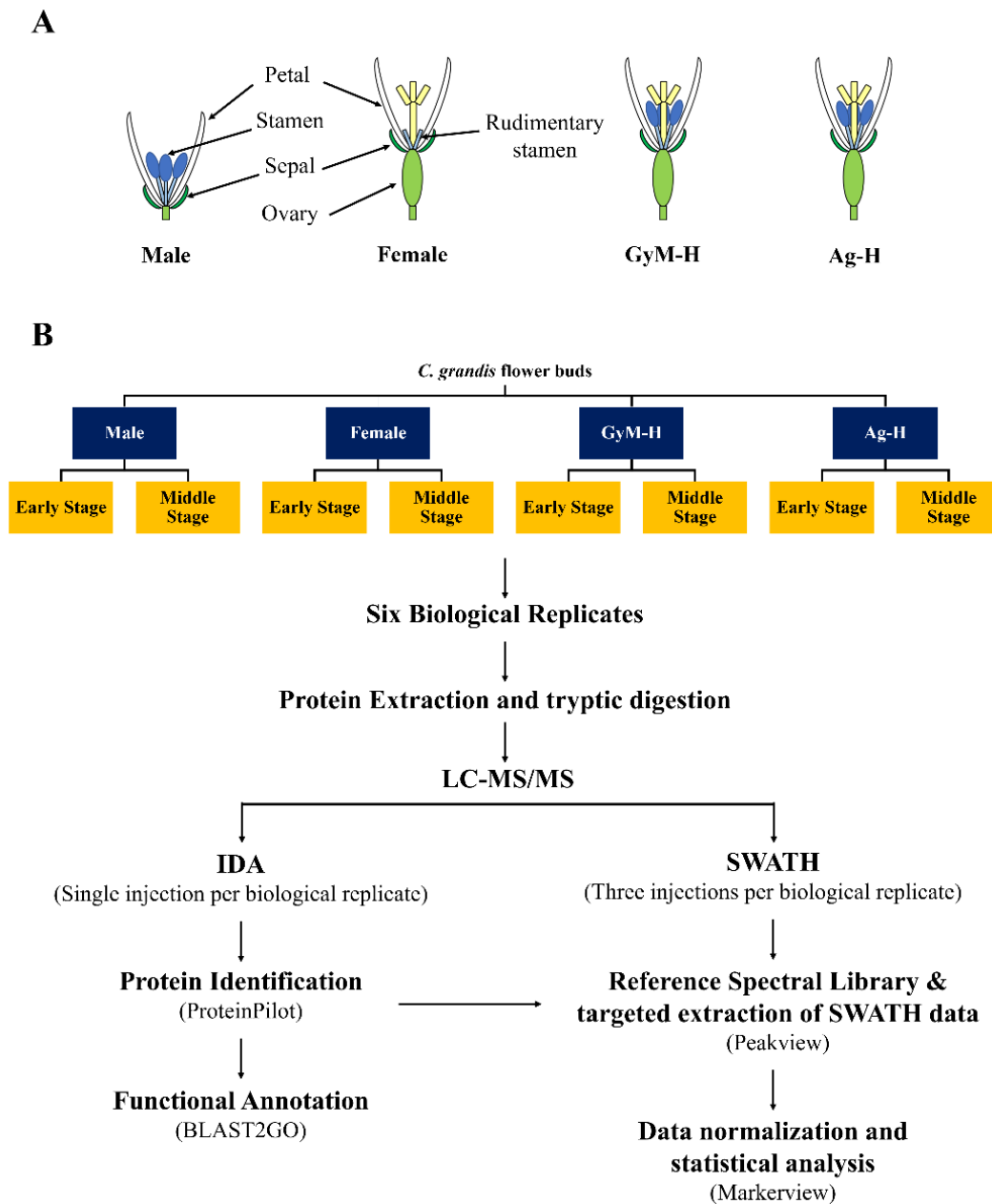
Figure 3-1. **Scheme for *Coccinia grandis* flower bud proteomics.** (A) Schematic representation of *C. grandis* flower types used in current study. (B) Experimental procedure for flower bud label-free proteomics.

3.2.2 Protein extraction

Protein extraction was carried out from flower buds of *C. grandis* as per Isaacson *et al.* (2006) with some modifications. Flower buds were crushed fine in liquid nitrogen and 50 mg of the

powdered samples, weighed in 15 mL tubes (TARSONS) were used for each biological replicate. Five mL of chilled 10 % TCA/Acetone with 2 % β-mercaptoethanol was added to these tubes, vortexed and incubated for at least 4 hours at -20 °C. The tubes were further centrifuged at 10000 rpm for 20 mins, and the supernatant was carefully discarded. The excess acetone in the pellet was evaporated in the fume-hood. Ten mg of insoluble PVPP (Polyvinylpolypyrrolidone) was added to this pellet followed by addition of 5 mL of extraction buffer containing 0.7 M sucrose, 0.1 M KCl, 0.5 M Tris-HCl, pH 7.5, 50 mM EDTA, 0.5 % SDS and β-mercaptoethanol at a final concentration of 2 % (v/v) and vortexed vigorously for 15 mins. The tubes were centrifuged for 20 mins at 4 °C and 10000 rpm. The supernatant was transferred to a fresh 15 mL tube to which 5 mL Tris equilibrated phenol (pH 8.0) was added and vortexed vigorously for 15 mins. Phase separation was carried out by centrifugation at 10000 rpm for 20 mins at 25 °C.

The upper phenolic phase was transferred to a fresh tube and washed with the extraction buffer (without SDS). The process was repeated twice, and the protein was precipitated using 3-5 volumes of chilled 0.1 M ammonium acetate in methanol containing 2 % β-mercaptoethanol. Pelleting and washing of the protein was performed as described previously (Isaacson *et al.*, 2006). The dried protein pellet was reconstituted in 0.1 % RapiGest SF (Waters) and quantitated using Bradford method (microtitre plate method, BioRad).

3.2.3 *In solution digestion and peptide purification*

For in-solution digestion of the extracted proteins, 100 µg of reconstituted proteins was taken in 1.5 mL LoBind tubes (Eppendorf) and diluted using 0.1 % RapiGest SF (prepared in fresh 50 mM ammonium bicarbonate solution) to make a final volume of 100 µL. This protein solution was heated at 80 °C for 15 mins on thermomixer at 1250 rpm. Five µL of 100 mM DTT (Dithiothreitol, freshly prepared) was added to the preheated protein solution and incubated at 60 °C, 1250 rpm for 15 mins. The protein solution was cooled to room temperature, and 5 µL of 200 mM iodoacetamide was added and incubated in dark for 30 mins at room temperature. Three µg of trypsin (Trypsin GOLD, Promega, MS-Grade) was added to the protein for digestion and incubated for 16 hours at 37° C and 1250 rpm. The reaction was stopped by adding 2 µL of formic acid (Sigma, MS-Grade) and incubated at 37 °C under static conditions for 45 mins. The tubes were centrifuged at 15000 rpm for 30 mins at 4 °C and supernatant was collected in a fresh 1.5 mL LoBind tube and stored at -80 °C until further use. The tryptic peptides were further

purified using C18 ZipTip (Millipore) by following manufacturer's instructions. The peptide fractions were vacuum dried and reconstituted using 3 % ACN with 0.1 % formic acid to yield a final concentration of 1µg/µL.

3.2.4 MS acquisition for IDA and SWATH

The reconstituted peptides were further diluted to yield concentrations of 3 µg/ 5µL for IDA acquisitions and 1 µg/ 5 µL for SWATH acquisitions. The LC separation methods and parameter settings for MS acquisitions have been adapted from Korwar *et al.* (2015).

LC Separation - Peptide digest were separated using Eksigent MicroLC 200 system (Eksigent, Dublin, CA) equipped with Eksigent C18-reverse phase column (100*0.3 mm, 3 µm, 120 Å). The sample was loaded onto the column with 97 % of mobile phase A (100% water, 0.1% FA) and 3 % of mobile phase B (100% ACN, 0.1% FA) at 8 µL/min flow rate. Peptides were eluted with a 120 min linear gradient of 3 to 50% mobile phase B. The column temperature was set to 40 °C and autosampler at 4 °C. The same chromatographic conditions were used for both IDA and SWATH acquisition.

Full MS/MS acquisition (IDA for creating SWATH library) - All samples were analyzed on AB-SCIEX 5600 Triple TOF mass-spectrometer in positive and high-sensitivity mode. The dual source parameters were optimized for better results: ion source gases GS1, GS2, curtain gas at 25 psi, temperature 200 °C and ion spray voltage floating (ISVF) at 5500 V. The accumulation time in full scan was 250 ms for a mass range of 350-1800 m/z. The parent ions are selected based on the following criteria's: ions in the MS scan with intensities more than 120 counts per second, charge stage between +2 to +5 and mass tolerance 50 mDa. Ions were fragmented in the collision cell using rolling collision energy (CE) with an additional CE spread of ± 15 eV.

Peptide spectral library and database search - IDA mass spectrometric files were searched using ProteinPilot software, version 5.0.1.0.4874 (AB SCIEX) with the Paragon algorithm against *in silico* translated in-house *C. grandis* flower bud transcriptome at 1 % FDR (false discovery rate). Similar analysis was also carried out using melon proteome and cucumber proteome as database. At the end, the ProteinPilot output file (.group) obtained using *in silico* translated in-house *C. grandis* flower bud transcriptome database was used as a standard peptide spectral library for downstream analyses.

SWATH MS - In SWATH-MS mode, the instrument was specifically tuned to optimize the quadrupole settings for the selection of precursor ion selection window 25 m/z wide. Using an isolation width of 26 m/z (containing 1 m/z for the window overlap), a set of 34 overlapping windows was constructed covering the precursor mass range of 400-1250 m/z. SWATH MS/MS spectra were collected from 100 to 2000 m/z. Ions were fragmented in the collision cell using rolling collision energy with an additional CE spread of ± 15 eV. An accumulation time (dwell time) of 96 ms was used for all fragmentation scans in high-sensitivity mode, and for each SWATH-MS cycle, a survey scan in high-resolution mode was acquired for 100 ms resulting in a duty cycle of 3.33 s. The source parameters were similar to that of IDA acquisition.

SWATH analysis was performed for six biological replicates and technical triplicates each from early-staged male, female and GyM-H buds as well as middle-staged male, female and GyM-H buds. The spectral alignment and targeted data extraction of SWATH-MS data was performed using Peakview software, version 2.2 (AB SCIEX). The peptide data (.MRKVW) files were used for quantification of proteins using Markerview software, version 1.3 (AB SCIEX). Normalization was performed using β-galactosidase and actin-7 (TRINITY_DN116897) peak area followed by the total area sum. The peptides with a $P \leq 0.05$ were considered for quantification. PCA-DA analysis was carried out using Markerview in order to study the relationship between all the samples and replicates followed by t-test to identify differentially expressed proteins (DEPs) among each pairwise comparisons in which fold change was set as $\geq$ 1.5, and significance of $P \leq 0.05$ was chosen as cut-off.

3.2.5 Protein-Protein Interaction Network analysis

Accession ID of *Arabidopsis thaliana* homologs of the proteins identified by proteomic analysis were used to study the protein-protein interaction using Cytoscape 3.6 (Shannon *et al.*, 2003). The STRING database for Arabidopsis thaliana was used to import the annotation, interaction and further analysis was performed using STRING App 1.2.2 (Szklarczyk *et al.*, 2017). Gene enrichment analysis was performed and proteins grouped under reproductive process and floral development were used for generating network representation.

3.2.6 Validation of differentially expressed proteins by qRT-PCR

For expression analysis of selected genes, qRT-PCR was carried out using aliquots of the same RNA samples that were used for RNA sequencing. Two micrograms of total RNA was used for complementary DNA (cDNA) synthesis by SuperScript IV reverse transcriptase (Invitrogen) using an oligo(dT) primer. *CgACT2* gene was used as reference gene for normalization. qRT-PCR was performed on CFX96 machine (BIO-RAD) with gene-specific forward and reverse primers. The reactions were carried out using Takara SYBR Premix Ex *Taq* II (Takara Bio Inc.) and incubated at 95 °C for 3 min, followed by 40 cycles of 95 °C for 15 s, 58 °C for 15 s and 72 °C for 15 s. PCR specificity was checked by melting curve analysis, and data were analysed using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

3.2.7 Primers used in this study

Primer Name	Sequence (5′ -> 3′)
EFE qF	CTTGGTGGAGAAAGAGGCGGAG
EFE qR	ACCTTGGCTCCTTCGCTTGAAA
AMS qF	CCGAACATCCGAAATCGACTGAAC
AMS qR	TAGTCAGTTTCGGTCGGTTTTTAGC
AIM1 qF	TCCGAAAATGCAGCGTCGAGAT
AIM1 qR	AGCAAACCATTTCAAAGCCAGCTC
PDIA6 qF	ATAAGTCGAGCTCCGCAGACTA
PDIA6 qR	TGTAAAGATCGGGAAATCGGGCT
TKPR1 qF	ACAAGGCCATCTTTCTTCACCTT
TKPR1 qR	AACTTCCACGTGTACGGTCGAG
VIP5 qF	ATGACTTCTGTGCCTTGCCCTC
VIP5 qR	CACAAATCCGGTCAGATCCGCT

3.3 Results

3.3.1 Flower bud proteome of C. grandis

The tryptic digested protein samples were analysed by tandem LC-MS in IDA (Information dependent acquisition) model, and the acquired data was processed by Paragon TM as per experimental scheme shown in Figure 3-1. A total of 1193777 spectra (68.8 % of total spectra) corresponding to 66842 distinct peptides representing total 3387 proteins (Additional File 1) were identified from different categories of *C. grandis* flower buds with FDR ≤ 1 (Unused ProtScore Cutoff > 2.0) using *in silico* translated in-house *C. grandis* flower bud transcriptome (Table 3-1). Similar analyses using melon proteome and cucumber proteome resulted in identification of 2381 and 2429 proteins. However, downstream analyses were carried out with 3387 proteins that were identified using *in silico* translated in-house *C. grandis* flower bud transcriptome database as it resulted in identification of higher number of proteins.

Table 3-1. Summary of protein identification with ABSCIEX Triple TOF platform using Paragon Algorithm with ProteinPilot v5.0.1.

Unused (Conf) Cutoff	Proteins Detected	Proteins Before Grouping	Distinct Peptides	Spectra Identified	% Total Spectra
>2.0 (99)	3387	14816	66842	1193777	68.8
>1.3 (95)	3855	17469	68548	1205044	69.5
>0.47 (66)	4649	21946	71945	1217799	70.2
Cutoff Applied: >0.05 (10%)	8979	59421	87507	1256111	72.4

3.3.2 Annotation of identified proteins

Proteins identified using ProteinPilot software were searched against viridiplantae subset of nr (non-redundant proteins) database using BLAST with an e-value threshold of 1e-3 (Figure 3-2A). Total 3377 proteins fetched at least one BLAST hit. More than 95 % of the proteins gave top BLAST hit against *Cucumis melo* (melon) or *Cucumis sativus* (cucumber). BLAST2GO was used to annotate and classify the identified proteins into three major gene ontology categories

(molecular function, cellular component and biological process) (Additional File 2). For biological process, majority of the proteins were involved in metabolic process (1956 proteins) and cellular process (1821 proteins) (Figure 3-2B).

Chapter 3: Proteomics approach for identification of sex-biased proteins in *C. grandis*
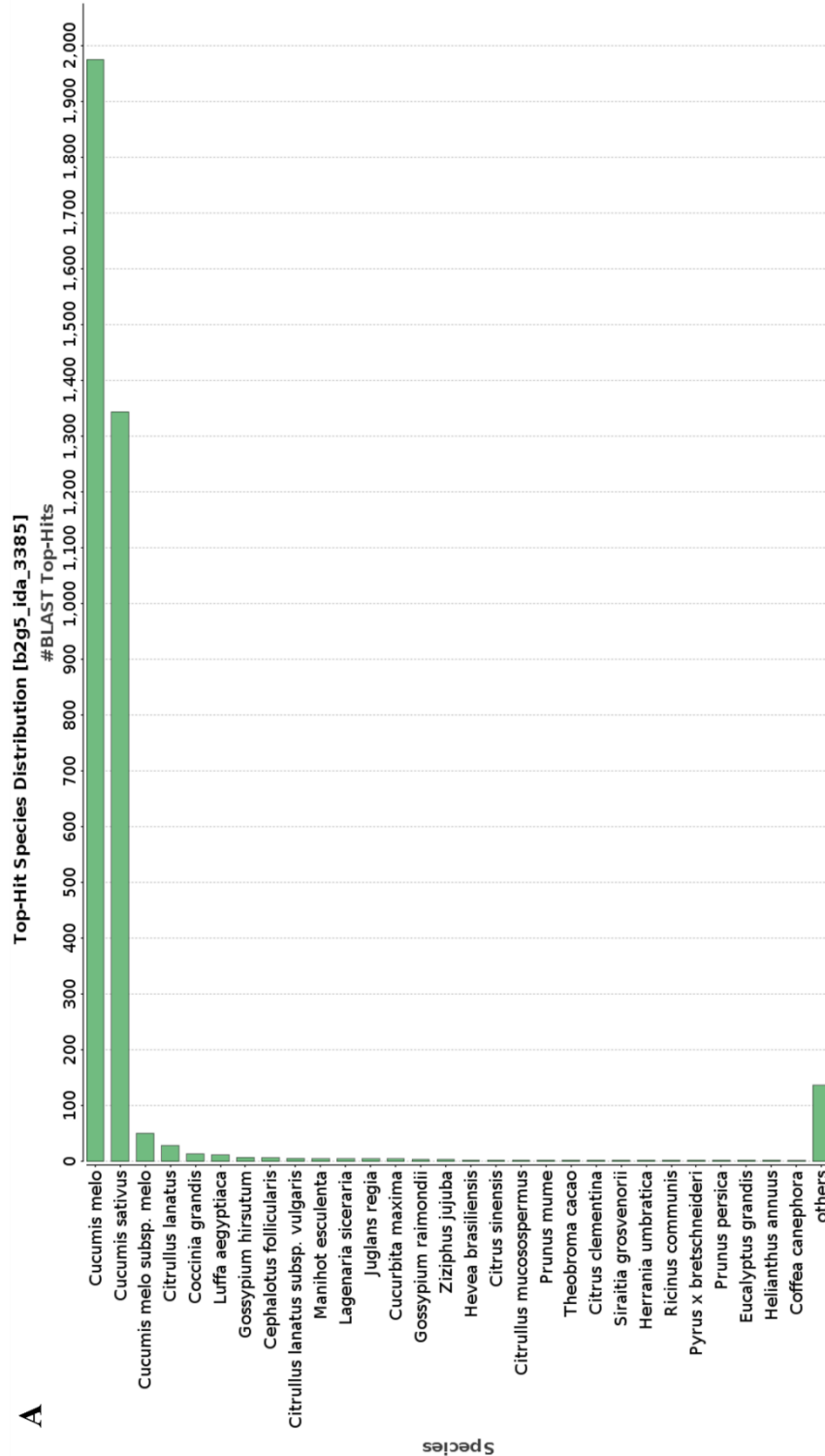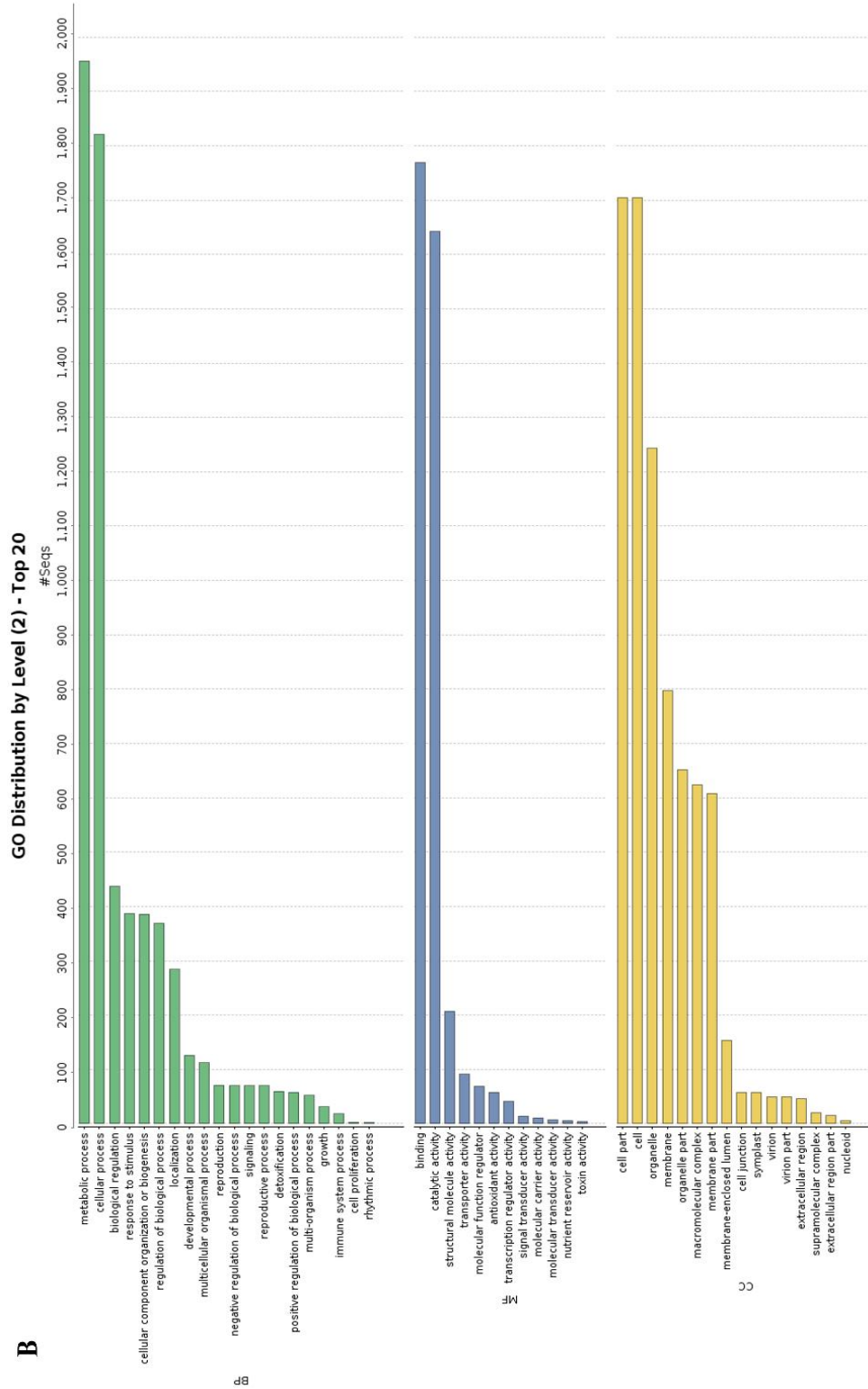
Figure 3-2. **Annotation of *C. grandis* flower bud proteome with BLAST2GO v5.** (A) BLAST Top-Hit species distribution when compared with nr database using tblastn (B) GO category (level 2) distribution of *C. grandis* flower bud proteome.

Chapter 3: Proteomics approach for identification of sex-biased proteins in *C. grandis*

Interestingly, we were able to detect 71 proteins annotated with reproduction (GO:0000003) term (Figure 3-2B). Binding (1769 proteins) and catalytic activity (1642 proteins) were two major categories for molecular function (Figure 3-2B). Also, 41 proteins with transcription regulator activity were detected (Figure 3-2B). Among cellular component, 1243 proteins associated with organelle, while 797 were related with membrane (Figure 3-2B) function. The most common enzymes were hydrolases, transferases and oxidoreductases in the identified proteome based on enzyme code distribution analysis (Figure 3-3).
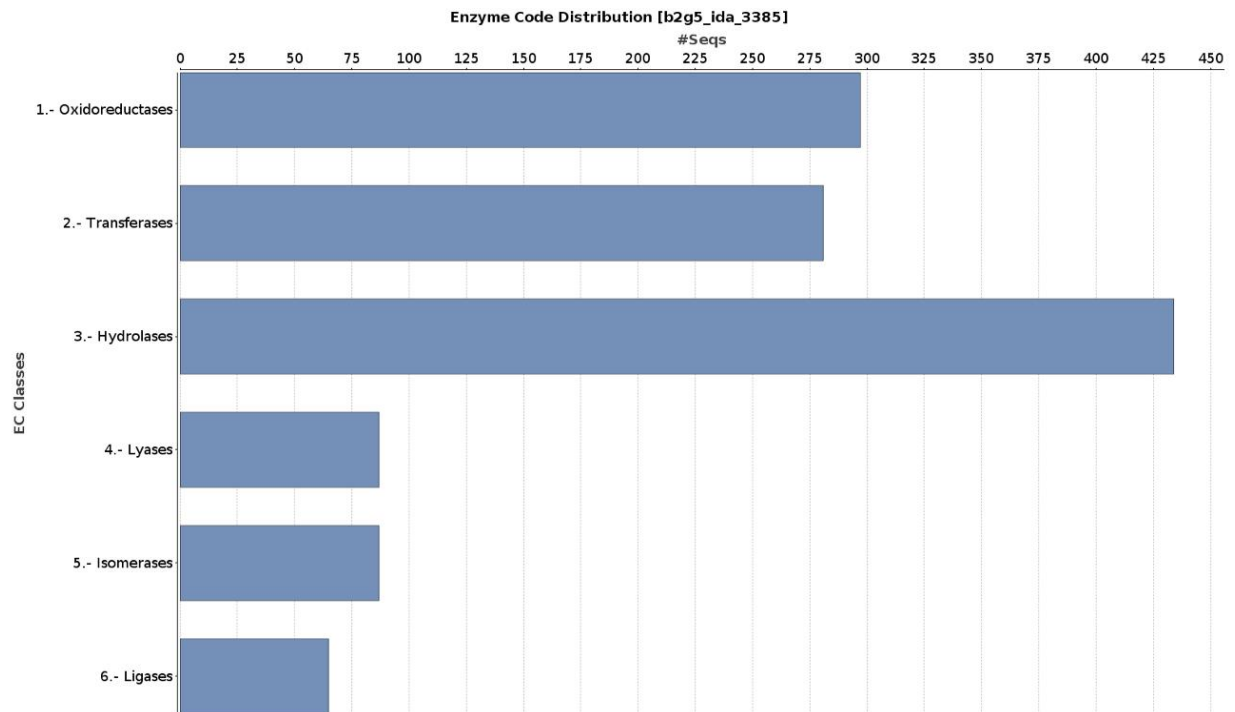
Figure 3-3. **Enzyme code distribution analysis using for the detected *C. grandis* flower bud proteins using BLAST2GO v5.**

3.3.3 *Differentially expressed proteins involved in flower development*

In order to understand how the expression of proteins was regulated during the development of male flower in contrast to female flower, quantitative proteomics was performed using label-free SWATH-MS to identify the differentially expressed proteins (DEPs). Based on these identified proteins from all the samples, SWATH-MS analysis led to the detection of total 2262 proteins that were quantified from 144 SWATH data files. PCA-DA analysis showed that replicates belonging to each category of the samples were clustered together. Additionally, upon PCA-DA

analysis, the samples were clearly separated as per their tissue sources or developmental stages (Figure 3-4).
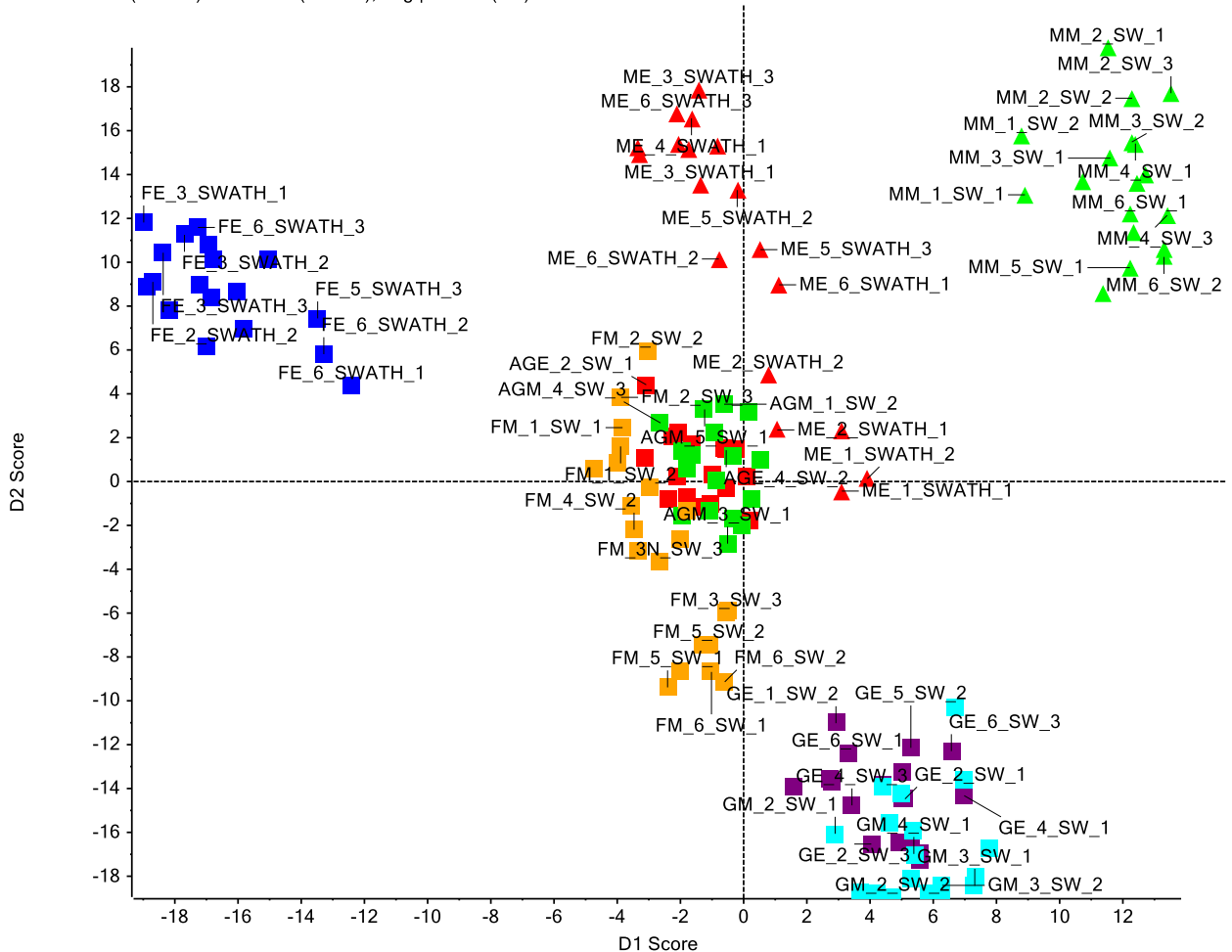
Figure 3-4. **PCA-DA analysis showing relationship between all the *C. grandis* flower bud samples as well as replicates in the study.** AGE (Red square), Early-staged Ag-H; FE (Blue square), Early-staged Female; GE (Purple square), Early-staged GyM-H; GM (Turquoise square), Middle-staged GyM-H; ME (Red triangle), Early-staged Male; MM (Green triangle), Middle-staged Male; AGM (Green square), Middle-staged Ag-H; FM (Orange square), Middle-staged Female.

This clearly reflects the significant variations in identified proteins as well as their expression levels among different tissues or stages. All pairwise comparisons were made to identify DEPs at fold change ≥ 1.5 and $P \leq 0.05$. Pairwise comparisons between the proteome profiles of different flower buds at early and middle stages of development are summarized in Table 3-2. Female_early vs male_early comparison was made in order to identify key players regulating stamen arrest during female bud development. Interestingly, proteins involved in

ethylene biosynthesis such as UBA2A (UBP1-associated protein 2A) (1.61X), UBA2C (UBP1-associated protein 2C) (2.1X) and EFE (Ethylene-forming enzyme) (1.64X) were upregulated in early-staged female buds compared to male buds (Figure 3-5A, B). Also, ERS (ethylene response sensor) (4.83X Middle) was enriched in middle-staged female buds compared to male buds (Figure 3-5B) (Additional File 3). We observed that the expression of AMS (ABORTED MICROSPORES) (0.46X) was downregulated in middle-staged female buds compared to male buds (Figure 3-6A, B & Figure 3-7) (Additional File 3).

Similarly, to identify the effect of AgNO$_3$ treatment on female plant that leads to stamen development, we compared protein profiles of Ag-H_early vs female_early buds. This comparison resulted in identification of many male function related proteins that were upregulated upon AgNO$_3$ treatment during early stage such as UGP2 (UTP-glucose-1-phosphate uridylyltransferase 2) (2.41X), EDA9 (EMBRYO SAC DEVELOPMENT ARREST 9) (2.05X), TKPR1 (Tetraketide alpha-pyrone reductase 1) (2.72X), C4H (Cinnamic acid 4-hydroxylase) (3.53X), TPLATE (1.56X), CDC2 (Cell division control protein 2 homolog A) (3.52), ANN5 (Annexin D5) (1.88X), RAB1C (Ras-related protein RABD2c) (3.13X), OASA1 (O-acetylserine sulfhydrylase 1) (3.56X), AtkdsA1 (Putative Aldolase-type TIM barrel family protein) (4.05X), TCTP (Translationally-controlled tumor protein) (2.52X) and ACAT2 (Acetoacetyl-CoA thiolase 2) (2.71X) (Figure 3-6B & Figure 3-7) (Additional File 3). However, many other male function-proteins like USP (UDP-sugar pyrophosphorylase) (1.69X), RPN10 (Regulatory particle non-ATPase 10) (1.39X) and LAP5 (LESS ADHESIVE POLLEN 5) (1.69X) were enriched in middle-staged Ag-H buds compared to female buds (Figure 3-6B) (Additional File 3). The other two comparisons (GyM-H_middle vs male_middle and Ag-H_middle vs male_middle) led to the identification of potential targets that may be involved in pollen germination and pollen tube growth. Interestingly, AMS (ABORTED MICROSPORES) (0.24X GyM-H, 0.14X Ag-H), UNE5 (UNFERTILIZED EMBRYO SAC 5) (0.47X Ag-H), iPGAM2 (0.49X GyM-H, 0.46X Ag-H), ACOS5 (Acyl-CoA Synthetase) (0.1X GyM-H, 0.03X Ag-H) and QRT3 (QUARTET 3) (0.14X GyM-H, 0.11X Ag-H) were highly expressed in middle-staged male buds compared to GyM-H and Ag-H buds (Figure 3-6B) (Additional File 3). Similarly, FUM1 (Fumarate hydratase 1) (0.68X GyM-H, 0.65X Ag-H) was enriched in early-staged male buds compared to GyM-H and Ag-H buds (Figure 3-6B) (Additional File 3).

Table 3-2. Summary of differentially expressed proteins (DEPs) in pairwise comparisons between male, female, GyM-H and Ag-H flower buds at early and middle stages of development.

Pairwise comparisons	Purpose	Number of DEPs
Ag-H Early vs. Female Early	To identify the effect of $AgNO_3$ treatment on female plants that leads to stamen development	1101
Ag-H Early vs. GyM-H Early	To identify common DEPs involved in stamen arrest compared to female buds	609
Ag-H Early vs. Male Early	Potential targets that may be involved in pollen development	870
Ag-H Middle vs. Female Middle	To identify DEPs involved in Ag^+-mediated stamen growth	329
Ag-H Middle vs. GyM-H Middle	To identify common DEPs involved in stamen arrest compared to female buds	707
Ag-H Middle vs. Male Middle	Potential targets that may be involved in pollen germination and pollen tube growth	801
Female Early vs. GyM-H Early	To identify DEPs involved in stamen arrest	1114
Female Early vs. Male Early	To identify proteins regulating stamen arrest during female bud development	644
Female Middle vs. GyM-H Middle	To identify DEPs involved in stamen arrest and pollen development	649
Female Middle vs. Male Middle	To identify DEPs involved in stamen arrest and pollen development	808
GyM-H Early vs. Male Early	Potential targets that may be involved in pollen development	885

Chapter 3: Proteomics approach for identification of sex-biased proteins in *C. grandis*
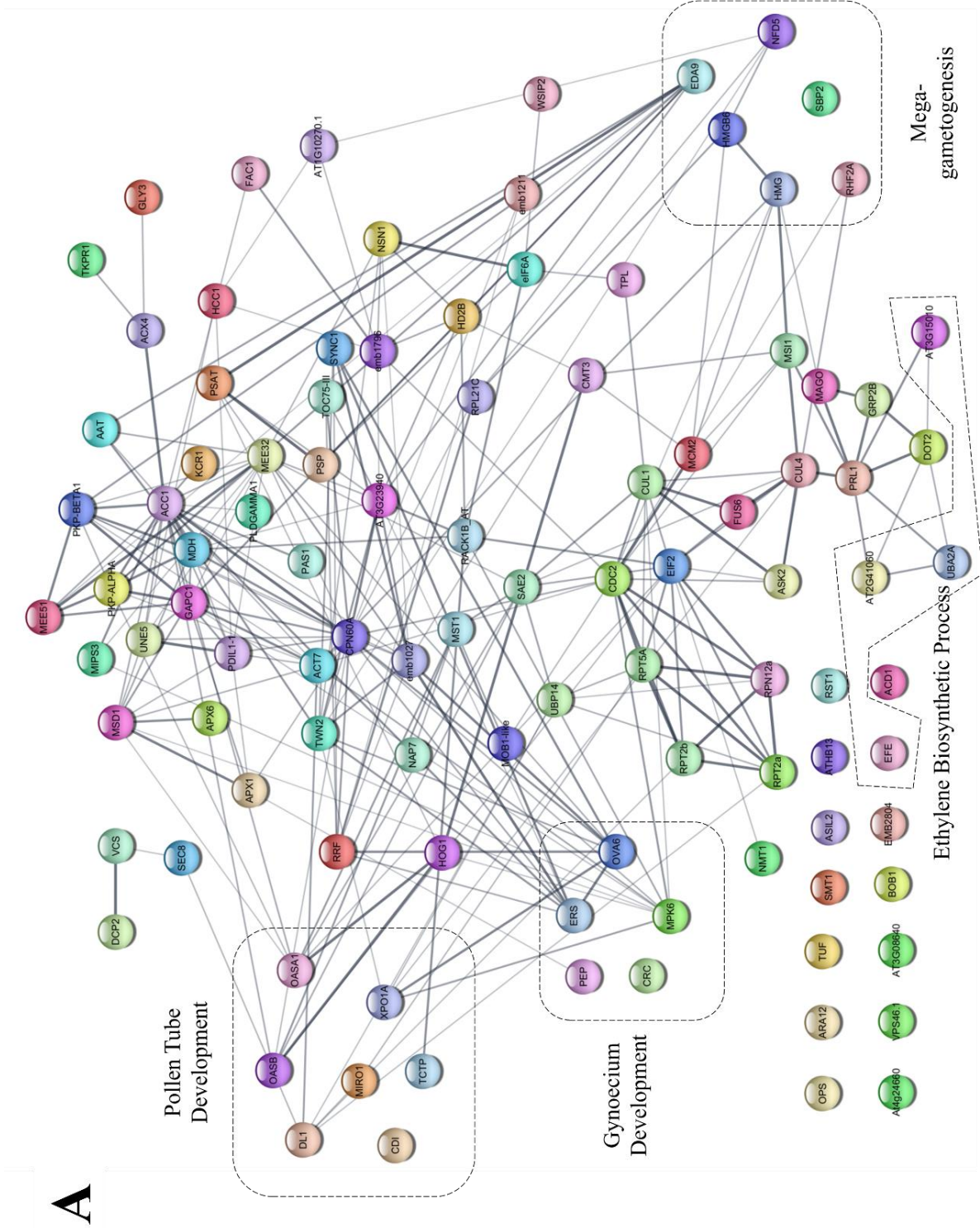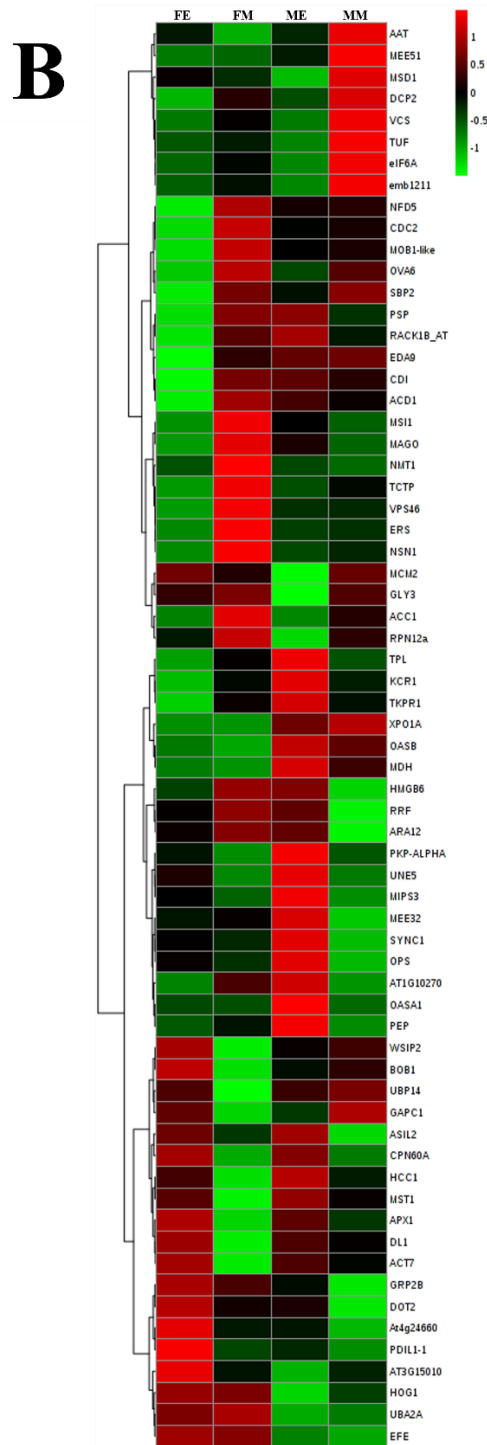
Figure 3-5. **Functional interactions and abundance profiles of proteins involved in female reproductive organ, embryo and seed development.** (A) Protein-protein functional interaction network generated using STRING. (B) Hierarchical clustering heatmap showing expression profiles for the selected proteins. Thickness of the line in interaction maps indicate the degree of confidence for prediction of the interaction. Dotted boxes with labels highlight the subset of proteins involved in respective processes.

Chapter 3: Proteomics approach for identification of sex-biased proteins in *C. grandis*

Chapter 3: Proteomics approach for identification of sex-biased proteins in *C. grandis*
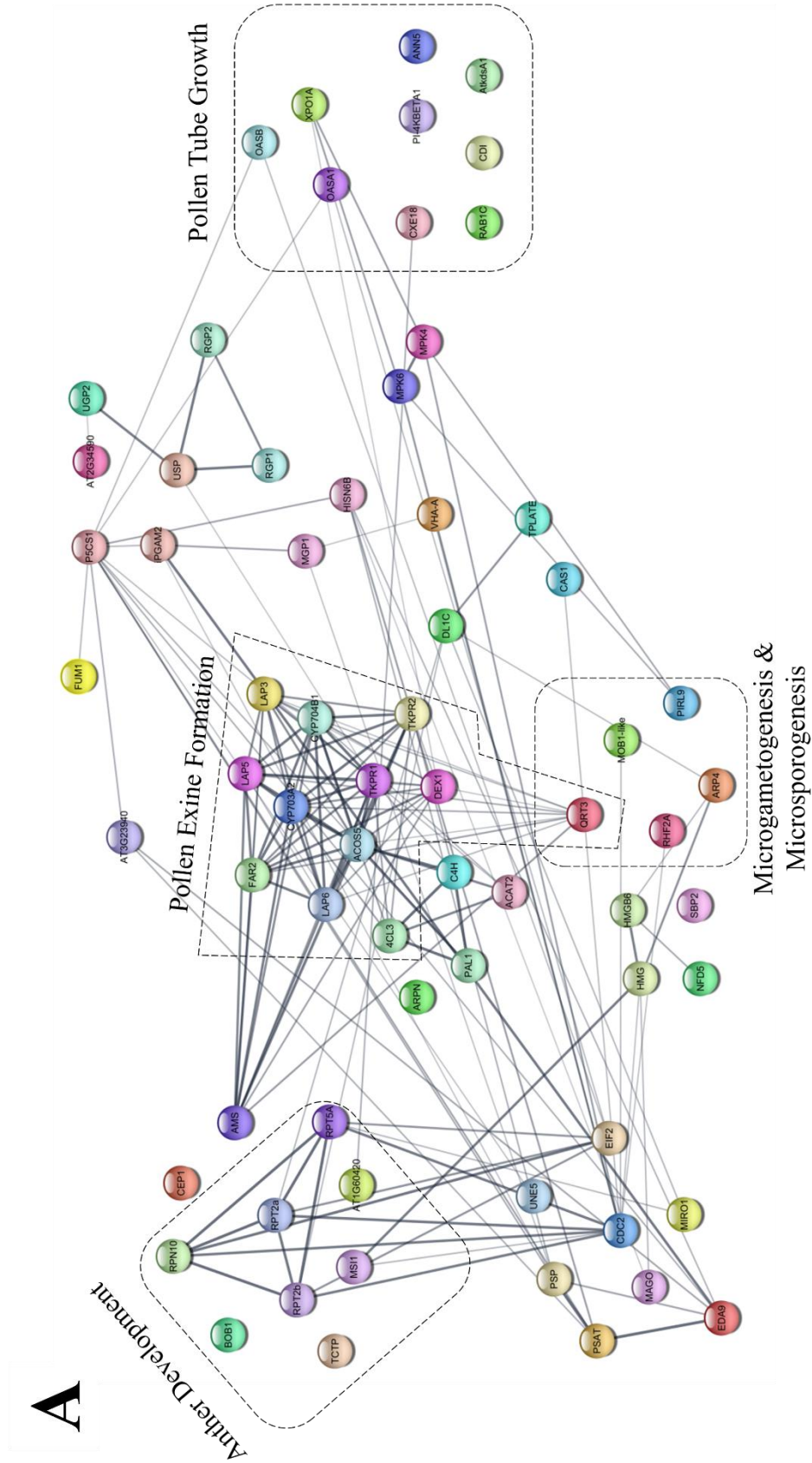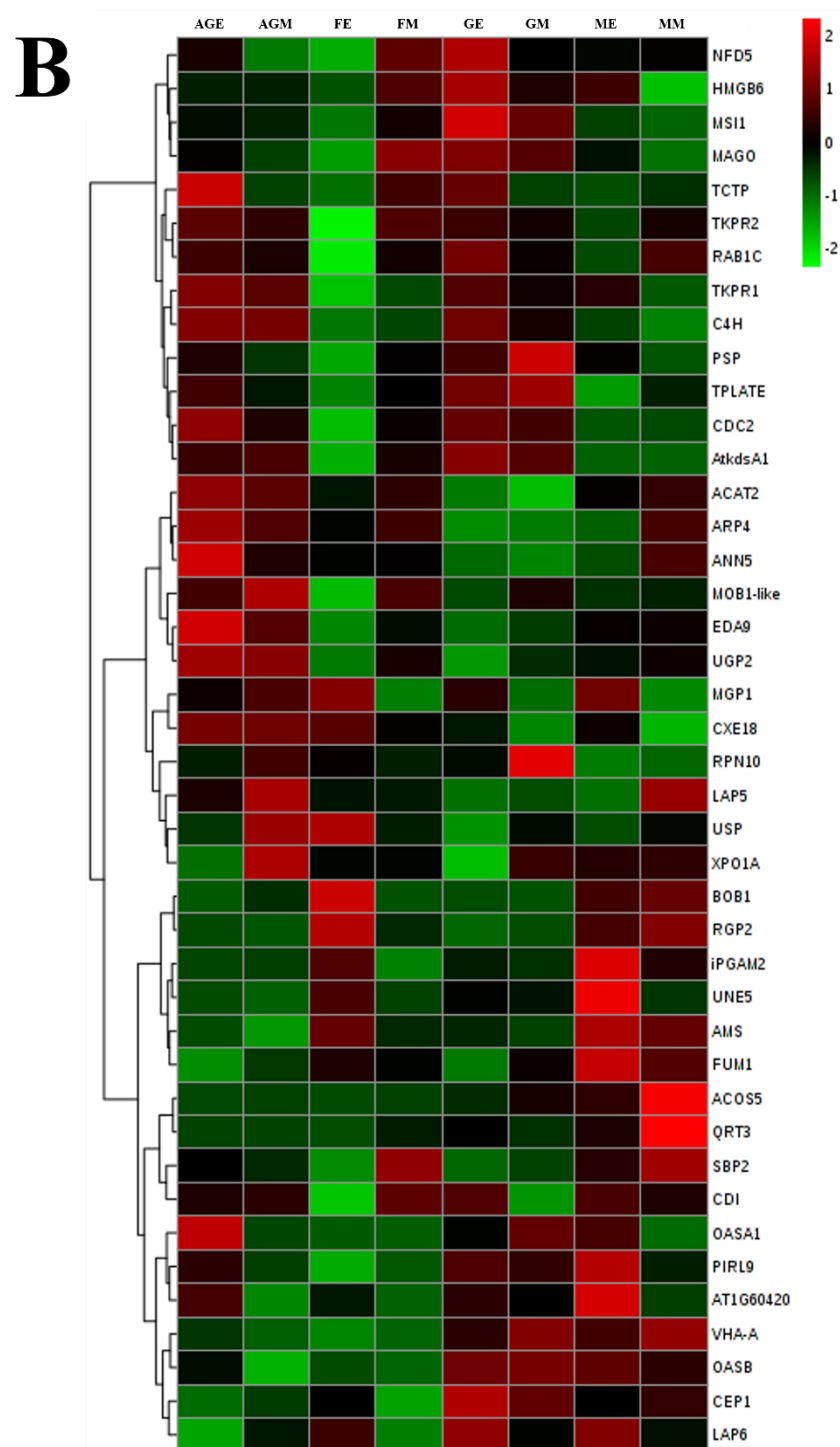
Figure 3-6. **Functional interactions and abundance profiles of proteins involved in male reproductive organ development.** (A) Protein-protein functional interaction network generated using STRING. (B) Hierarchical clustering heatmap showing expression profiles for the selected proteins. Thickness of the line in interaction maps indicate the degree of confidence for prediction of the interaction. Dotted boxes with labels highlight the subset of proteins involved in respective processes.

Chapter 3: Proteomics approach for identification of sex-biased proteins in *C. grandis*

3.3.4 Transcript abundance corroborates changes in protein profiles

Transcript levels of the selected proteins were validated by qRT-PCR analysis using the same tissues. *Coccinia* homologs of *AIM1*, *EFE*, *VIP5*, *AMS*, *TKPR1* and *PDIA6,* showed interesting transcript abundance profiles with respect to flower bud types (Figure 3-7). *EFE* showed two-fold expression in early-staged female buds as compared to male buds. Pollen development related transcripts *AMS*, *TKPR1* and *PDIA6* were enriched in male, GyM-H and Ag-H buds compared to female buds. *TKPR1* showed more than 100-fold expression in middle-staged male buds compared to female buds and was enriched in middle-staged GyM-H and Ag-H buds. Candidate gene *AMS* exhibited a 100-fold expression in early-staged male buds compared to female buds; >5-fold levels compared to GyM-H buds and 10-fold increase compared to Ag-H buds. However, there were marginal differences between the transcript and the protein levels in some of the genes. For example, PDIA6 protein levels were highest in early-staged male buds, while Ag-H buds showed very low levels. However, in contrast to protein abundance, *PDIA6* transcript was most enriched in Ag-H buds.

3.3.5 Protein interaction network reveals presence of potential candidate genes from reproductive organ development

The proteins identified in relation to reproductive organ, embryo and seed development were further used for protein-protein functional interaction mapping. STRING analysis for 42 male reproductive organ development proteins resulted in a functional interaction map indicating anther and pollen development, pollen tube growth as well as microgametogenesis and microsporogenesis related functions (Figure 3-6A) (Additional File 4). Expression analysis of these proteins showed that majority were enriched in male, GyM-H and Ag-H buds compared to female buds (Figure 3-6B). Similarly, another STRING analysis of 66 identified proteins from female reproductive organ, embryo and seed development led to the identification of interaction networks governing gynoecium development, ethylene biosynthesis, megagametogenesis and pollen tube development (Figure 3-5A) (Additional File 4). Importantly, proteins from gynoecium development and ethylene biosynthesis related clusters were enriched in female buds (Figure 3-5B). This might lead to the stamen inhibition in female buds and explain the absence of gynoecium in male buds.

A

B
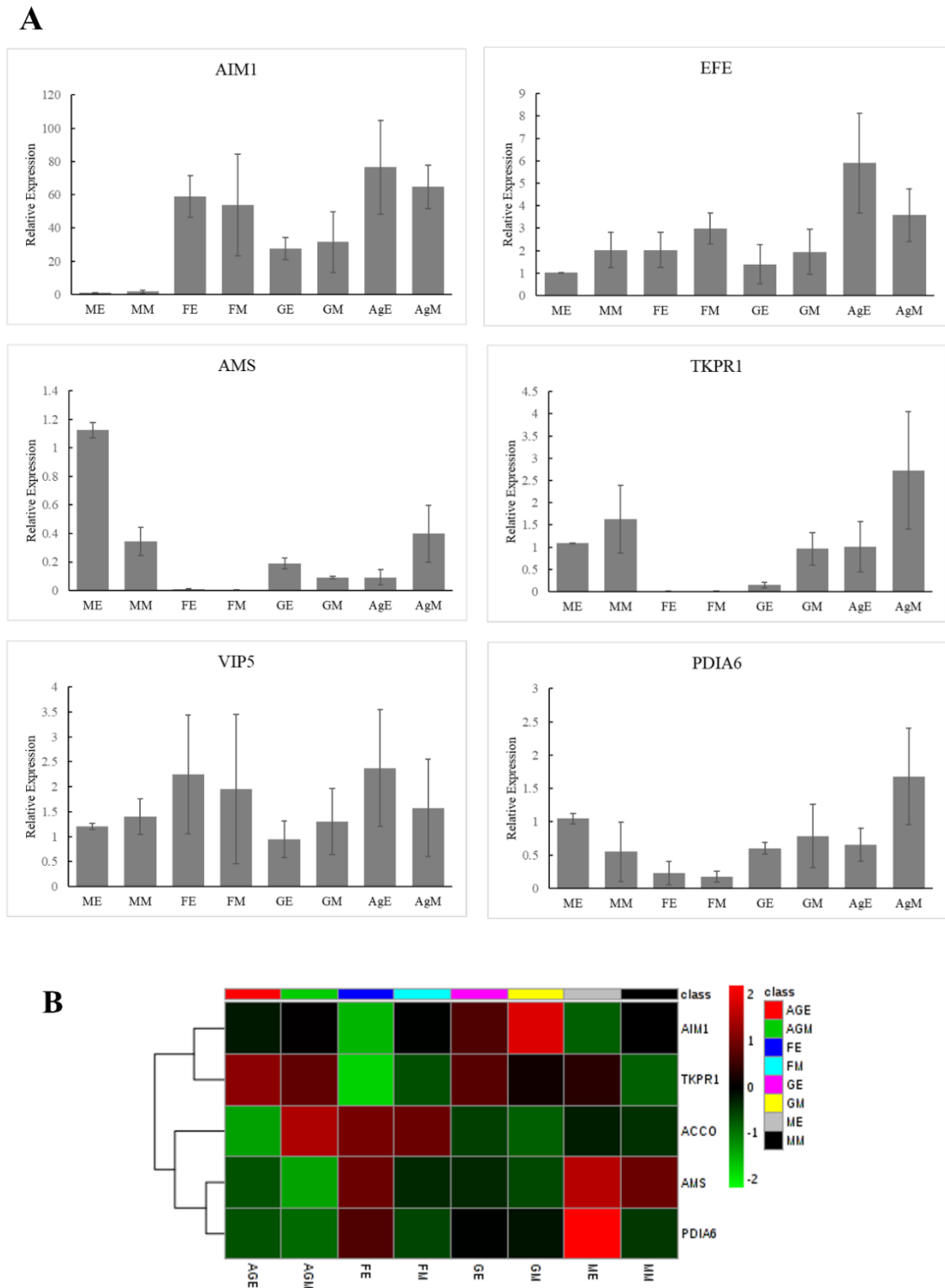
Figure 3-7. **Validation of selected DE proteins by qRT-PCR with three biological replicates.** (A) The relative expression in the sample of early-staged male (ME) was set to 1 for plotting the qRT-PCR data. (B) Hierarchical clustering heatmap for the protein expression levels as quantified by SWATH-MS. VIP5 protein was detected in the IDA study, but it was not quantified by SWATH-MS. AgEA, Early-staged

Chapter 3: Proteomics approach for identification of sex-biased proteins in *C. grandis*

Ag-H A; AgEB, Early-staged Ag-H B; FEA, Early-staged Female A; FEB, Early-staged Female B; GEA, Early-staged GyM-H A; GEB, Early-staged GyM-H B; GMA, Middle-staged GyM-H A; GMB, Middle-staged GyM-H B; MEA, Early-staged Male A; MEB, Early-staged Male B; MMA, Middle-staged Male A; MMB, Middle-staged Male B; AgMA, Middle-staged Ag-H A, AgMB, Middle-staged Ag-H B; FMA, Middle-staged Female A; FMB, Middle-staged Female B.

3.4 Discussion

3.4.1 Comparative profiles of proteins involved in female reproductive functions

Our study is one of the few proteomic reports focused on identification of proteins involved in unisexual flower development. We compared the protein profiles of male and female buds of *Coccinia* at early as well as middle stages of development and identified proteins related to male and female reproductive organ development. At early stage, female flower buds of *C. grandis* exhibit the growth of both stamens and carpels. However, before the flowers reaches to middle stage of development, stamen gets arrested in female buds. The mechanism of stamen arrest has been well studied in monoecious relatives of *Coccinia* such as melon, cucumber and watermelon, however, we have no clue about the stamen arrest in unisexual flower of *Coccinia* (Boualem *et al.*, 2008, Boualem *et al.*, 2009, Boualem *et al.*, 2015, Boualem *et al.*, 2016). ACS (ACC synthase), an enzyme involved in ethylene biosynthesis is found to be expressed at very high levels in female buds of melon, cucumber and watermelon compared to the male buds. High levels of ethylene in female bud act as an inhibitor of stamen development (Boualem *et al.*, 2008, Boualem *et al.*, 2009, Switzenberg *et al.*, 2014, Boualem *et al.*, 2015, Boualem *et al.*, 2016). However, it is not known if the ACS function is conserved in dioecious cucurbits. Previously, we have shown that GO terms related to ethylene signalling were enriched in female buds of *C. grandis* (Devani *et al.*, 2017). In our current study, however, we couldn't detect ACS from the flower buds. Instead, we found that the inducers of ACS, UBA2A (UBP1-associated protein 2A) and UBA2C (UBP1-associated protein 2C) had very high expression in female buds compared to male buds of *C. grandis* (Figure 3-5A, B). Also, we found that at transcript level, *CgACS* was 10-fold enriched in female buds compared to male buds of *C. grandis* at early stage (Figure 2-4E). Similar high *CgACS* transcript levels in female *C. grandis* buds has also been reported by Mohanty *et al.* (2017). In Arabidopsis, Kim *et al.* (2008) showed that overexpression of UBA2A and UBA2C induces the expression of various ACS genes. EFE (Ethylene-forming enzyme), another enzyme involved in ethylene biosynthesis was also enriched in early as well as middle-

staged female buds compared to male buds (Figure 3-5A, B & Figure 3-7). *CsACO2* gene, an EFE homolog is known to be essential for carpel development in cucumber (Chen *et al.*, 2016). Recently, Tao *et al.* (2018) showed that Ethylene responsive factor (ERF110) mediates ethylene-regulated transcription of sex determination related orthologous gene *ACS11* in melon and cucumber. This evidence further supports the role of ethylene in governing the sex expression in cucurbits. Moreover, we observed that female buds of *C. grandis* also showed much higher accumulation of ERS (Ethylene response sensor), an ethylene receptor at middle stages of development compared to male buds (O'Malley *et al.*, 2005) (Figure 3-5A, B). This indicates that ethylene levels as well as ethylene responses are high even in female buds of dioecious *C. grandis* similar to the female buds of monoecious cucurbits. This perhaps explains the cause of stamen arrest in *C. grandis*. Further, functional studies are required to validate this hypothesis. Apart from this, OVA6 (OVULE ABORTION 6) involved in ovule development was enriched in female buds as the male flowers of *C. grandis* don't show any sign of female reproductive organs (Berg *et al.*, 2005) (Figure 3-5A, B). Levels of proteins involved in embryo development such MAGO, VPS46 (Vacuolar protein-sorting-associated protein 46), PDIL1-1(Protein disulfide isomerase-like 1-1), GRP2B (Glycine-rich protein 2b), At4g24660 (Zinc-finger homeodomain protein 2) were also higher in female buds compared to male counterparts (Pagnussat *et al.*, 2005, Andème Ondzighi *et al.*, 2008, Spitzer *et al.*, 2009, Yang and Karlson, 2011) (Figure 3-5A, B). MAGO, a protein involved in pollen tube guidance was expressed in female buds at middle stage of development indicating the maturation of female buds for supporting fertilization (Johnson *et al.*, 2004) (Figure 3-5A, B).

3.4.2 *AgNO₃ treatment induces changes in protein profiles affecting sex modification*

In our previous study, we have shown that foliar spray of AgNO$_3$ at optimal concentrations (35mM) can induce stamen development in newly emerging flower buds of female plants resulting in sex modification (Ghadge *et al.*, 2014). The morphologically hermaphrodite flowers (Ag-H) have full-sized stamens surrounding the carpels. Based on our *de novo* transcriptomics study and the literature available for sex determination mechanism in monoecious relatives of *Coccinia* (melon, cucumber and watermelon), we hypothesized that Ag$^+$ ions bring about the sex modification by inhibiting ethylene signalling (Saito *et al.*, 2007, Boualem *et al.*, 2008, Boualem *et al.*, 2009, Li *et al.*, 2009, Boualem *et al.*, 2015, Boualem *et al.*, 2016, Chen *et al.*, 2016,

Devani *et al.*, 2017). *CmACS7* acts as inhibitors of stamen development in melon by playing a role in ethylene biosynthesis (Boualem *et al.*, 2008). In our transcriptomic study, we showed that AgNO$_3$ spray resulted in inhibition of ethylene signalling in *C. grandis* (Devani et al., 2017). In the present report, we observed that *Coccinia* homologs of many genes involved in pollen development such as USP (UDP-sugar pyrophosphorylase), RPN10 (Regulatory particle non-ATPase 10), UGP2 (UTP-glucose-1-phosphate uridylyltransferase 2), EDA9 (EMBRYO SAC DEVELOPMENT ARREST 9), TKPR1 (Tetraketide alpha-pyrone reductase 1), C4H (Cinnamic acid 4-hydroxylase), TPLATE, CDC2 (Cell division control protein 2 homolog A), ANN5 (Annexin D5), RAB1C (Ras-related protein RABD2c) were enriched in Ag-H buds compared to female buds (Smalle *et al.*, 2003, Van Damme *et al.*, 2006, Kotake *et al.*, 2007, Schilmiller *et al.*, 2009, Grienenberger *et al.*, 2010, Park *et al.*, 2010, Dobritsa *et al.*, 2011, Peng *et al.*, 2011, Toujani *et al.*, 2013, Yang *et al.*, 2014, Zhu *et al.*, 2014) (Figure 3-6A, B). LAP5 (LESS ADHESIVE POLLEN 5), TKPR1 and TKPR2 (Tetraketide alpha-pyrone reductase 2) which are involved in sporopollenin biosynthesis and pollen exine formation were upregulated upon AgNO$_3$ treatment (Dobritsa *et al.*, 2010, Grienenberger *et al.*, 2010, Kim *et al.*, 2010, Dobritsa *et al.*, 2011) (Figure 3-6A, B & Figure 3-7). Also, *Coccinia* homologs for OASA1 (O-acetylserine sulfhydrylase 1), AtkdsA1 (Putative Aldolase-type TIM barrel family protein), TCTP (Translationally-controlled tumor protein), ACAT2 (Acetoacetyl-CoA thiolase 2), ANN5, RAB1C proteins involved in pollen germination and pollen tube growth were enriched in Ag-H buds compared to female buds (Berkowitz *et al.*, 2008, Delmas *et al.*, 2008, Peng *et al.*, 2011, Birke *et al.*, 2013, Zhu *et al.*, 2014)(Figure 3-6A, B). Although these proteins were expressed upon AgNO$_3$ treatment, the pollens of Ag-H buds failed to germinate *in vitro* as well as did not fertilize the female buds (Ghadge *et al.*, 2014). Interestingly, Ag$^+$ ions were able to induce the development of stamens even in absence of Y-chromosome, which would otherwise get arrested in female plants (Ghadge *et al.*, 2014). We speculate that Y-chromosome may be essential for providing pollen fertility as the pollens from Ag-H flowers buds were sterile in nature.

3.4.3 Pollen maturation in male, GyM-H and Ag-H buds

Comparison of protein accumulation profiles for male buds to those of GyM-H and Ag-H buds resulted in identification of many proteins that were expressed at different levels in these flower types. *Coccinia* homologs of proteins such as LAP6 (LESS ADHESIVE POLLEN 6), CEP1

(KDEL-tailed cysteine endopeptidase), OASB (O-acetylserine sulfhydrylase B), VHA-A (V-type proton ATPase subunit a3) and PIRL9 (Plant intracellular Ras-group-related LRR protein 9) were enriched in male and GyM-H buds compared to Ag-H and female buds (Figure 3-6A, B). LAP6 protein known to be involved in pollen exine formation / sporopollenin biosynthesis in Arabidopsis, was found to be enriched in male and GyM-H buds compared to Ag-H and female buds of *C. grandis* (Dobritsa *et al.*, 2010, Kim *et al.*, 2010) (Figure 3-6A, B). In contrast, LAP5, (also involved in pollen exine formation) was seen to be enriched in male and Ag-H buds but not in GyM-H and female buds (Dobritsa *et al.*, 2010, Kim *et al.*, 2010) (Figure 3-6A, B). A mutation in *Arabidopsis* CEP1, a KDEL-tailed cysteine endopeptidase, has been shown to exhibit aborted tapetal PCD and decreased pollen fertility with abnormal pollen exine (Zhang *et al.*, 2014b). On the other hand, VHA-A (V-type proton ATPase catalytic subunit A) protein is strictly required for proper development of male gametophyte (Dettmer *et al.*, 2005). Consistently, VHA-A had a high expression in male and GyM-H buds (Figure 3-6A, B). Plant intracellular ras-group-related LRR protein (PIRL9) shown to play an important role in the process of microgametogenesis and pollen development (Forsthoefel and Vernon, 2011). PIRL9 had high expression in male and GyM flower buds (Figure 3-6A, B). OASB (O-acetylserine sulfhydrylase) required for pollen tube elongation and fertilization also showed similar expression profile (Birke *et al.*, 2013) (Figure 3-6A, B).

Many proteins however, were enriched specifically in male buds with very low levels in GyM-H, Ag-H and female buds. AMS (ABORTED MICROSPORES), UNE5 (UNFERTILIZED EMBRYO SAC 5) and iPGAM2 *Coccinia* homologs were enriched in middle-staged male buds (Figure 3-6A, B). UNE5 is known to be important in the process of pollen tube development (Boavida *et al.*, 2009). Mutation in transcription factor AMS (ABORTED MICROSPORES) results in defective tapetum development, abnormal microspores and non-viable pollens (Xu *et al.*, 2010). Similarly, *Coccinia* homologs ACOS5 (Acyl-CoA Synthetase) and QRT3 (QUARTET 3) involved in pollen exine formation were also enriched in middle-staged male buds. (Rhee *et al.*, 2003, de Azevedo Souza *et al.*, 2009, Dobritsa *et al.*, 2011) (Figure 3-6A, B). Another protein FUM1 (Fumarate hydratase 1), involved in pollen tube development was enriched in early-staged male buds compared to GyM-H and Ag-H buds (Boavida *et al.*, 2009) (Figure 3-6A, B). All these observations indicate that these proteins probably govern the pollen viability and lack of their expression in GyM-H and Ag-H buds leads to the sterile pollens.

3.5 Conclusion

Current study has led to the identification of many proteins expressed in the flower buds of male, female and gynomonoecious forms of *Coccinia grandis* at early and middle stages of development. Also, the changes in proteome profile upon AgNO$_3$-mediated sex modification in female plants were revealed. We observed that proteins involved in ethylene biosynthesis such as UBA2A, UBA2C and EFE were upregulated in female buds of *C. grandis* indicating high levels of ethylene similar to the female buds of monoecious cucurbits such as melon, cucumber and watermelon. This could suggest that the role of ethylene in stamen inhibition might be conserved not only in monoecious cucurbits but also in dioecious *C. grandis.* AgNO$_3$ treatment was able to induce pollen development in female plants of *C. grandis*. However, the pollens from Ag-H and GyM-H buds were sterile in nature. There are two possibilities that could explain the cause of pollen sterility in GyM-H and Ag-H flower buds. The flower buds of hermaphrodites and males could differ in the timing or level of gene expression leading towards the incomplete development of pollens. Alternatively, genes governing pollen fertility may be Y-linked and absent from GyM-H and Ag-H forms. As till now, complete genome sequence data of *Coccinia* is not available; it would be difficult to ascertain this. Future investigation is required to unravel the mechanistic link between ethylene-mediated stamen inhibition as well as the role of Y-chromosome in pollen viability. Overall, our current proteome dataset will be helpful in identification of key molecular players involved in the development of unisexual flowers in Cucurbitaceae, the second largest horticultural family in terms of economic importance.

3.6 Supporting information

Additional File 1. Detailed report of protein identification from *C. grandis* flower buds using Paragon algorithm with ProteinPilot v5.0.1. Available at https://figshare.com/s/9d20d3eed5c9cea36077

Additional File 2. BLAST2GO annotation table for the *C. grandis* flower bud proteins detected in this study. Available at https://figshare.com/s/9d20d3eed5c9cea36077

Additional File 3. Normalized peak areas for each of the quantified proteins across all the samples and pairwise differential expression analysis results with respective fold-changes as well as *P*-values. Available at https://figshare.com/s/ab4b7b16835496340aa7

Additional File 4. Normalized peak areas for the proteins involved in male and female reproductive organ development, seed development and embryo development along with their respective *Arabidopsis* accession IDs. Available at https://figshare.com/s/b4dd1e5a607449a76c0c

Results presented in this chapter have been communicated to *The Plant Journal,* and the manuscript is currently undergoing review.

Devani R.S., Chrimade T., Sinha S., Bendahmane A., Banerjee A.K.[*] and Banerjee J.[*] (2018). Flower bud proteome reveals modulation of sex-biased proteins potentially associated with sex expression and modification in dioecious *Coccinia grandis*. **Under review.**

Chapter 4: Development of Virus Induced Gene Silencing (VIGS) tool for functional characterization of genes in *Coccinia grandis*

4.1 Background

Through our transcriptomics and proteomics approach, we have generated detailed information regarding the sex-biased genes for *Coccinia grandis*. However, functional understanding of these genes would be really difficult due to the lack of regeneration and transformation protocol in *C. grandis*. Our attempts for *in vitro* regeneration and transformation of *C. grandis* were unsuccessful. Virus-induced gene silencing (VIGS) provides a powerful tool to enable the gene function analysis through knock-down for such species that are not amenable to stable genetic transformation (Becker and Lange, Baulcombe, 1999). VIGS approach has never been reported earlier for *Coccinia grandis*. There is only one report till now for induction of VIGS in the species belonging to Cucurbitaceae family (melon, cucumber, etc.) (Igarashi *et al.*, 2009). So, the major challenge for us was to select an ideal viral system which can infect and penetrate the floral meristem of *C. grandis*.

4.1.1 Molecular mechanism of virus-induced gene silencing (VIGS)

Virus-induced gene silencing technique takes the advantage of plant's sequence-specific defence mechanism against viruses known as post-transcriptional gene silencing (PTGS) (Figure 4-1). Double-stranded RNA (dsRNAs) derived from DNA-based viral vectors using host RNA-dependent RNA polymerase (RdRp) can act as inducers of RNA-interference (RNAi) mediated gene silencing (Covey *et al.*, 1997, Al-Kaff *et al.*, 1998). In case of RNA viruses, replication intermediates generated using viral RdRPs also act as inducers of RNAi (RNA-interference) (Ratcliff *et al.*, 1997, Donaire *et al.*, 2008). DICER-like enzymes then degrade these dsRNAs into short interfering RNA (siRNAs) duplexes, which are approximately 21–24 nucleotides in length. These siRNA molecules act as guide and provide sequence-specificity for the degradation of complementary viral RNAs by the RNA-induced silencing complex (RISC) (Waterhouse and Fusaro, 2006, Ding and Voinnet, 2007). PTGS against the inducing virus ultimately results in viral recovery conferring *de novo* resistance against viruses closely related to the primary inoculum (Ratcliff *et al.*, 1997).

In order to take the advantage of the PTGS defence mechanism induced upon viral infection to achieve target gene knock-down, several RNA and DNA viruses have been modified to develop VIGS vectors (Baulcombe, 1999, Senthil-Kumar and Mysore, 2014). The target gene to be knocked-down can be cloned inside the modified infectious viral DNA (DNA virus) or cDNA

(RNA virus) containing VIGS vectors (Figure 4-1). Different methods can be employed for the introduction of the viral vectors in the plant, for example, mechanical inoculation (Igarashi *et al.*, 2009), *Agrobacterium*-mediated infiltration (Senthil-Kumar and Mysore, 2014) and biolistic inoculation (Sasaki *et al.*, 2011). siRNAs generated by PTGS will be directed against both the viral genome in the VIGS vector and host transcript that has sequence similarity to the inserted DNA fragment. This will eventually lead to the degradation of the target transcript and hence knock-down of the corresponding target gene. The virus-derived silencing signal is further amplified and spreads systemically throughout the plant (Kalantidis *et al.*, 2008).
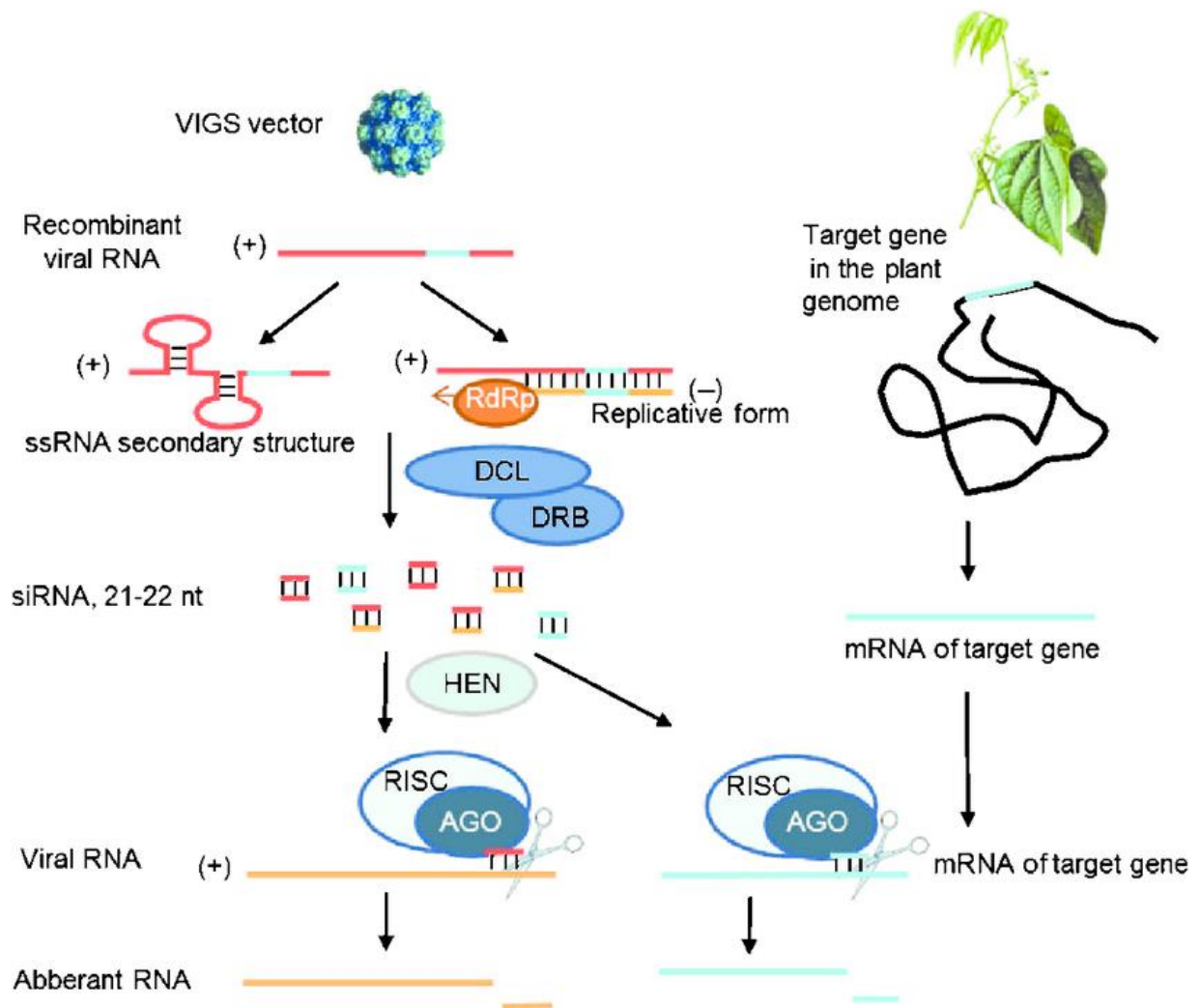
Figure 4-1. **Molecular mechanism of virus-induced gene silencing (VIGS) induced by RNA viruses in plants.** Abbreviations: AGO, Argonaute protein; DCL, Dicer-like protein; DRB, dsRNA-binding protein; HEN, Hua enhancer protein; RdRp, RNA dependent RNA polymerase; RISC, RNA-induced silencing complex; siRNA, short-interfering RNA; ssRNA, single-stranded RNA. Reproduced from Pflieger *et al.* (2013).

4.1.2 Commonly used Viral vectors for VIGS induction

Several virus vectors have been developed for inducing VIGS in plants (Senthil-Kumar and Mysore, 2014), including vectors of tobacco mosaic virus (TMV) (Kumagai *et al.*, 1995), potato virus X (PVX) (Chapman *et al.*, 1992, Ruiz *et al.*, 1998), tomato golden mosaic virus (TGMV) (Peele *et al.*, 2001), and tomato yellow leaf curl China virus satellite DNA (Tao and Zhou, 2008). Tobacco rattle virus (TRV)-derived vectors have been used for inducing VIGS in plants belonging to the Solanaceae family including tomato, potato, *N. benthamiana*, and also in the Brassicaceae model plant *Arabidopsis thaliana* (Liu *et al.*, 2002, Brigneti *et al.*, 2004, Burch-Smith *et al.*, 2004, Fu *et al.*, 2005, Cai *et al.*, 2006). Modified Apple latent spherical virus-derived vectors have also been reported to infect and induce VIGS in a broad range of plants including those belonging to the Brassicaceae, Leguminaceae, Solanaceae and Cucurbitaceae families (Igarashi *et al.*, 2009). Several other viral vectors have been used in different plants species, a detail list of which is mentioned in the report of Senthil-Kumar and Mysore (2014).

4.1.3 Advantages and limitations of virus induced gene silencing (VIGS) as a tool for fast forward genetic studies in plants

The VIGS technique has several advantages over other functional genomics approaches (Burch-Smith *et al.*, 2004, Senthil-Kumar and Mysore, 2011). Firstly, VIGS can be used to generate knock-down without the need to generate stable transformants, and it can be used for studying a loss-of-function phenotype for a specific gene within a single generation. Secondly, the VIGS technique is less time consuming, and the silencing phenotype can be observed within 15 to 20 days. Another important advantage of the VIGS approach is that the problem of redundancy can be solved by using a target sequence from the highly conserved region of a gene family. Also, knowledge about partial gene sequence is sufficient for VIGS based knock-down approaches. Unlike stable transgenics, VIGS can be used to study the function of genes, whose mutation leads to lethal phenotypes. Thus, VIGS offers an easy method for determining the functions of the genes in a short time, and it can also be utilized in high throughput functional genomics in plants (Dong *et al.*, 2007).

Apart from the advantages listed above, VIGS has its own set of limitations too (Senthil-Kumar and Mysore, 2011). Sometimes, it becomes difficult to dissect out the cause of phenotype as apart from the target gene knock-down, viral infection may also alter the plant metabolism

(Burch-Smith *et al.*, 2004, Zhang *et al.*, 2010, WU *et al.*, 2011). There are reports that showed that target gene inserted in VIGS vectors might interfere with the viral replication in some cases and hence, many viruses are known to delete the target gene insert upon replication (Bruun-Rasmussen *et al.*, 2007, Cakir and Tör, 2010). Also, the gene silencing achieved by VIGS is transient, incomplete and non-uniform throughout the plant. Most of the viruses are excluded from actively growing apical tissues and hence cannot be used for studying the function of genes involved in meristematic functions (Burch-Smith *et al.*, 2004). Also, the infectivity of virus varies even with different genotypes of the same species and may need specific standardization in many cases. Hence, the advantages and disadvantages should be weighed carefully with respect to the goal of the study before choosing VIGS tool for gene knock-down.

There were two major objectives for our study:

(i)   To establish and standardize an efficient virus-induced gene silencing technique for gene function analysis in *C. grandis.*

(ii)   To study the role of putative candidate genes potentially involved in sex determination and differentiation (such as *CgACS*, *CgPI,* etc.) in *C. grandis* using the VIGS technique.

In this study, we have reported an efficient protocol for inducing VIGS for the silencing of endogenous genes in *Coccinia grandis* using the apple latent spherical virus (ALSV) vector by the mechanical inoculation method. Attempts are being made to study the function of *CgACS* and *CgPI* in *C. grandis* using the standardized tool.

4.2 Materials and methods

4.2.1 *Explant collection and surface sterilization for in vitro regeneration and transformation trials of C. grandis*

Initially, transformation and plant regeneration attempts were made with leaf and tendril explants of. *C. grandis*. Young leaves and tendrils were collected from the *C. grandis* plants grown at greenhouse facility IISER Pune. After washing the leaves and tendrils in running tap water for 5 min, 70 % (v/v) ethanol treatment was given for 1 min for surface sterilization. Following 3 washes in sterile distilled water, the leaves were treated with 1.0 % (v/v) sodium hypochlorite solution for 10 min with occasional agitation. Finally, the leaves and tendrils were washed for

five times with sterile distilled water. The surface sterilized leaves were then cut in approximately 1 cm^2 pieces such that midrib portion was present in the pieces. The surface sterilized tendrils were also cut into 1 cm long pieces.

4.2.2 Co-culture with Agrobacterium and callus induction

Leaf and tendril explants were pre-cultured on callus-induction media for four days. Callus-induction media for leaf explants consisted of Murashige and Skoog basal salt mixture with B5 vitamin (MSB5), 3 % sucrose supplemented with 6.6 μM 2,4-Dichlorophenoxyacetic acid (2,4-D), 3.3 μM 6-Benzylaminopurine (BAP) and 0.8 % agar (Thiruvengadam and Chung, 2011). Whereas for tendril explants, two different callus-induction media were used: (i) MS, 3 % sucrose supplemented with 2.2 μM 1-Naphthaleneacetic acid (NAA) and 6 μM Thidiazuron (TDZ), (ii) MS, 2 % Glucose supplemented with 2.2 mg/L Zeatin riboside (Zr) (Raju *et al.*, 2014). Meanwhile, *Agrobacterium tumefaciens* strain GV2260 harbouring pBI121 plasmid was streaked on Luria-Bertani (LB) agar plate containing 50 mg/L kanamycin and 10 mg/L rifampicin. A single colony was grown in 5 mL liquid LB media containing 50 mg/L kanamycin and 10 mg/L rifampicin antibiotics at 28 °C on a rotatory shaker (200 rpm). Three mL of this culture was transferred to 10 mL of fresh LB media devoid of any antibiotics and cultured until the OD$_{600}$ reached to 0.8-1.0. Pre-cultured tendrils and leaves (adaxial/top side down) were then placed in the liquid infection medium (IM) (Half-MS medium augmented with 1.5 % sucrose and 100 μM acetosyringone (3',5'-dimethoxy-4'-hydroxyacetophenone). 100 μL of *Agrobacterium* culture was added to the plate. The plate was subjected to gentle shaking at 25°C for 30 min. The leaves and tendrils were then blotted dry on sterile filter paper for 3 min to remove the excess infection medium. Following which they were transferred to a co-cultivation medium (Callus-induction media + 200 μM acetosyringone) and incubated for three days at 25 °C in darkness. After the co-cultivation, explants were washed with sterile distilled water for five times and blotted dry on sterile filter paper. Then the explants were maintained on callus-induction media supplemented with 50 mg/L Kanamycin (for selection of transformants) and 250 mg/L cefotaxime at 25 °C under long-day (16 h light, 8 h dark) conditions until callus growth was achieved. In order to assess the *Agrobacterium* infection, GUS staining of callus was carried out using 5-bromo- 4-chloro-3-indolyl β-D-glucuronide as per the method of Jefferson *et al.* (1987).

4.2.3 Regeneration of C. grandis plantlets from transformed callus

Calli originated from leaf explants were sub-cultured on shoot-induction media composed of MSB5, 3% sucrose augmented with 8.8 μM BAP, 2.2 μM 2,4-D, 50 mg/L kanamycin and 250 mg/L Cefotaxime (Thiruvengadam and Chung, 2011). Similarly, calli originated from tendril explants were sub-cultured on MS medium supplemented with 4.0 mg/L BAP for shoot induction (Raju *et al.*, 2014). However, all attempts were failed, and hence we moved to VIGS based knock-down system for *C. grandis*.

4.2.4 Construction of TRV vectors

pTRV1, pTRV2-MCS, and pTRV2-*NtPDS* vectors were procured from Arabidopsis Biological Resource Center (ABRC, https://abrc.osu.edu/order-stocks). pTRV1 and pTRV2-MCS vectors developed by Liu *et al.* (2002) were used in this study. pTRV2-MCS vector was cut using the restriction enzymes *EcoRI* and *XbaI*. *EcoRI* and *XbaI* restriction sites were added to the primers designed for amplification of sequences to be cloned. cDNA from wild-type *C. grandis* was used to amplify *CgPDS A* (224 bp), *CgPDS B* (154 bp), *CgPI A* (219 bp)*, CgPI B* (251 bp) and *CgPI C* (630 bp)*.* The PCR amplification was performed using *PfuTurbo* DNA polymerase. PCR products were digested using *EcoRI* and *XbaI* restriction enzymes. The digested PCR products were ligated to *EcoRI-XbaI* cut pTRV2 vector using T4 DNA Ligase (NEB) using manufacturer's protocol. The plasmids were then transformed into competent cells of *E. coli* DH5α. True transformants were selected using kanamycin resistance and cultured in 5 mL of LB medium at 37 °C overnight. The plasmid DNAs were isolated from the cultured cells by the alkaline lysis method. To identify the inserts, plasmid DNAs were restricted with *EcoRI-XbaI* followed by electrophoresis on a 1 % agarose gel. The plasmids having an insert with the desired length were then confirmed by DNA sequencing. Positive clones of pTRV1 and pTRV2-GOI (GOI- Gene of Interest) were transformed into competent *Agrobacterium* GV2260 cells.

4.2.5 Construction of ALSV vectors

pEALSR1 and pEALSR2L5R5 vectors were kindly provided by Prof. Yoshikawa, Iwate University, Japan (Igarashi *et al.*, 2009). A 300 bp fragment of the *CgPDS* gene and a 220 bp fragment of the *CgACS* gene were amplified from total RNA of *C. grandis* using a forward primer containing an *XhoI* site and a reverse primer containing a *BamHI* site. The ALSV vectors

containing the target gene was constructed using the pEALSR2L5R5 vector. The amplified cDNA fragments were each double digested with *XhoI* and *BamHI* and ligated to the pEALSR2L5R5 vector digested with same restriction enzymes. The plasmids were then transformed into competent cells of *E. coli* NEBTurbo. True transformants were selected using ampicillin resistance and cultured in 5 mL of LB medium at 37 °C overnight. The plasmid DNAs were isolated from the cultured cells by the alkaline lysis method. To identify the inserts, plasmid DNAs were restricted with *XhoI* and *BamHI* followed by electrophoresis on a 1 % agarose gel. The plasmids having an insert with the desired length were then confirmed by DNA sequencing. The plasmid containing the *CgPDS* gene fragment was designated as pEALSR2L5R5-*CgPDS,* and the plasmid containing the *CgACS* gene fragment was designated as pEALSR2L5R5-*CgACS*. pEALSR2L5R5 was used as the empty vector control. The plasmids pEALSR1, pEALSR2L5R5, pEALSR2L5R5-*CgPDS* and pEALSR2L5R5-*CgACS,* were propagated in *E. coli* NEBTurbo. The transformed *E. coli* cells were streaked on an LB agar medium containing 100 mg/mL ampicillin and allowed to grow for 12 hours at 37 °C. A primary culture of transformed *E. coli* was obtained by inoculating a single colony in 5 mL of LB medium containing 100 mg/mL of ampicillin, grown for 8 hours at 37 °C with vigorous shaking. 1 mL of the primary culture was inoculated in 100 mL of LB medium containing 100 mg/mL of ampicillin, to obtain a secondary culture of the transformed *E. coli* cells grown for 16 to 18 hours at 37 °C with vigorous shaking. Plasmid DNAs were purified using Promega PureYield™ Plasmid Midiprep System according to the manufacturer's protocols. The purified plasmids were dissolved in nuclease-free water at a concentration of 2 µg/µL and stored at -20 °C until use for inoculation.

4.2.6 *Construction of agroinfiltration compatible vectors for ALSV VIGS*

In order to improve the infection efficiency, simplify the infection process as well as reduce the time frame for gene silencing, agroinfiltration compatible ALSV in pCAMBIA-based binary vectors were constructed using the cloning strategy reported by Kon and Yoshikawa (2014). For constructing ALSV RNA1 vector, pCAMBIA1302 plasmid was used. Firstly, an ALSV RNA1 fragment (~ 0.5 kb) including 3′-cDNA, poly-A sequences and NOS-T was extracted from pEALSR1 vector using *Bam*HI-*Eco*RI digestion. This fragment was cloned into the multiple cloning site (MCS) of pCAMBIA1302 plasmid. Second fragment (~1.9 kb) containing the

CaMV35S promoter and the ALSV RNA1 5′-cDNA was released from pEALSR1 using *Hind*III-*Xba*I digestion. This fragment was subsequently cloned into the pCAMBIA1302 plasmid containing the first fragment. Finally, the third fragment (~6.2 kb) containing ALSV RNA1 cDNA was released from pEALSR1 by digesting it with *Sac*I-*Bam*HI. The third fragment was then cloned into pCAMBIA1302 plasmid containing the first and second fragments to generate pC1302ALSR1.

Similarly, pCAMBIA1301 plasmid was used to generate ALSV RNA2 vector. pCAMBIA1301 was digested with *Xho*I to remove the *hygromycin phosphotransferase* (*HPT*) gene and blunt-ended withT4 DNA polymerase. Firstly, a fragment of ALSV RNA 2 (~0.9 kb) containing 3′-cDNA, poly-A sequences as well as NOS-T was extracted from pEALSR2L5R5 vector using *Kpn*I-*Eco*RI digestion. This fragment was cloned into pCAMBIA1301 plasmid from which *HPT* gene was removed. Second fragment (~2.2 kb) containing *CaMV35S* promoter and the ALSV RNA2 5′-cDNA was released from pEALSR2L5R5 using *Hind*III-*Bam*HI digestion and cloned into pCAMBIA1301 plasmid containing the first fragment. Finally, the third fragment (~1.8 kb) containing ALSV RNA2 cDNA was released from pEALSR2L5R5 using *Bam*HI-*Spe*I digestion. This fragment was cloned into pCAMBIA1301 plasmid containing the first and second fragments to generate pC1301ALSR2.

4.2.7 Plant growth conditions and viral inoculation

(i)   TRV-based VIGS

Single *Agrobacterium* colonies of both pTRV1 and pTRV2-GOI (GOI- Gene of Interest) were inoculated separately in 2 mL LB liquid medium containing rifampicin (10 µg/mL) and kanamycin (50 µg/mL). Primary cultures were incubated at 28 °C in shaker at 200-250 rpm. for 10-14 h. 0.5 mL of this primary culture was subcultured into 10 mL LB liquid medium containing rifampicin (10 µg/mL) and kanamycin (50 µg/mL). Secondary culture was incubated at 28 °C in shaker at 200-250 r.p.m. for 5-6 h until OD_{600} (Optical density measured at 600 nm wavelength) of 0.5-0.6 is reached. Cultures were centrifuged at 3,500 g for 5 min at room temperature. Supernatant was discarded and 10 mL *Agrobacterium* induction buffer (10 mM MES, pH5.7,10 mM $MgCl_2$) containing 200 µM acetosyringone was used to resuspend the cells. Cultures were incubated at room temperature for 3 h in a shaker at ~50 r.p.m. Acetosyringone was used to induce *Agrobacterium* virulence (Vir) genes essential for delivery of T-DNA

carrying virus into the plant genome. Cultures were centrifuged at 3,500 g for 5 min, and cells were resuspended independently in ~5-7 mL induction buffer depending on the final OD required. OD_{600} ranging from 0.4 - 0.7 was used in this study. pTRV1 and pTRV2-GOI cultures were mixed in 1:1 (vol/vol) ratio. *C. grandis* seeds were germinated in greenhouse at 22 °C under long day conditions. Seedlings at cotyledonary stage were selected for viral inoculation. Infiltration of pTRV1+pTRV2-GOI culture into the cotyledons by needleless syringe (Syringe inoculation) was done after pricking a small injury with a needle. Minimum of 10 seedlings were infiltrated for each construct in every experiment. Plants were maintained at 20-22 °C (optimal temperature for VIGS) in a greenhouse at long day conditions (16 h light 8 h dark). Silencing phenotypes were checked starting from 10 days post infiltration. Phenotypes were noticed after ~ 2-3 weeks of inoculation. For molecular validation of gene silencing, tissues were collected in liquid nitrogen after 3 weeks and stored in -80 °C for future use.

(ii)   ALSV based VIGS

Seeds of *Chenopodium quinoa* were sown in pots and grown in a greenhouse under long day conditions (16 h light and 8 h dark) at 22 °C till the plants reached six-true leaf stage. The plasmids pEALSR2L5R5, pEALSR2L5R5-*CgPDS* and pEALSR2L5R5-*CgACS,* were mixed with the plasmid pEALSR1 separately at a concentration of 2 µg/µL each. *C. quinoa* plants at six-true leaf stage were used for infection. The inoculation of infectious cDNA clones was conducted on four bottom leaves per plant. Prior to inoculation, carborundum was spread on the leaves to cause injury to the plant. A mixture of pEALSR1 and pEALSR2L5R5 / pEALSR2L5R5-*CgPDS* / pEALSR2L5R5-*CgACS* (10 µL) was dropped onto the surface of each leaf and mechanically inoculated on it by swabbing it on the leaf surface using a finger covered by rubber gloves. The plants were incubated for 5 minutes to allow the infectious plasmids to enter the plant body through the injuries. After 5 minutes, the surface of inoculated leaves was thoroughly washed using distilled water to remove excess carborundum. The symptoms of ALSV infection consisting of chlorotic spots developed in the upper un-inoculated leaves after 14-20 days post infection. The leaves showing symptoms were homogenized in 2 volumes of inoculation buffer [0.1 M Tris-HCl (pH 7.5), 0.1 M NaCl, 0.05 M $MgCl_2.6H_2O$]. Using this extract, a secondary infection was done on *C. quinoa* (6 leaf stage) in order to increase the viral titre. Leaves showing severe symptoms were collected and stored at -80 °C until use for

inoculation. Leaves of *C. quinoa* (stored at −80 °C) infected with the ALSV vectors were homogenized in 2 volumes of inoculation buffer. *C. grandis* seedlings at the cotyledonary stage were used for inoculation. Prior to inoculation, carborundum was spread on the surface of cotyledons to cause injury. 30 µL of the leaf extract was dropped onto the surface of each cotyledon and mechanically inoculated on it by swabbing it on the leaf surface using a finger covered by rubber gloves. The plants were incubated for 5 minutes to allow the infectious viral particles to enter the plant body through the injuries. After 5 minutes, the surface of inoculated cotyledons was thoroughly washed using distilled water to remove excess carborundum. The inoculated plants were covered with plastic bags in order to maintain high moisture conditions and were maintained under greenhouse conditions.

4.2.8 RNA extraction and qRT-PCR analysis

Inoculated plants were scored for viral infection approximately 20 days post inoculation. Leaf tissues from six plants each of untreated, TRV-*CgPDS* A, TRV-*CgPDS* B, TRV-*CgPI* C, ALSV-*CgPDS*-infected and ALSV-*CgACS*-infected *C. grandis* were collected and crushed in a motor pastel using liquid nitrogen. Approximately 100 mg of tissue was used for total RNA extraction using Trizol reagent (Takara, Bio Inc.). 1 µg of the extracted RNA was reverse transcribed to cDNA using MMLV Reverse Transcriptase (Promega), using polyT primers. Samples were PCR screened to detect the presence of TRV RNA using primers TRV Chk F and TRV Chk R. Whereas, R2ALS-F and R2ALS-R primers were used for detection of ALSV-RNA2 in plants subjected to ALSV VIGS. The PCR was set using *Taq* DNA polymerase with the following sequential protocol: 94 ℃ for 5 minutes followed by 35 cycles of 94 ℃ for 10 seconds, 52 ℃ for 10 sec and 74 ℃ for 10 seconds followed by final extension at 74 ℃ for 5 minutes. The PCR products was separated on a 2 % agarose gel containing 0.5 mg/L ethidium bromide. DNA band intensities were imaged using G:Box Gel Documentation System (Syngene).

The relative levels of *CgPDS* transcript expression for all un-infected and ALSV-*CgPDS* infected *C. grandis* plants (n=6) were determined using quantitative RT-PCR (qRT-PCR). *CgACT2* gene was used as reference gene for normalization. A 10 µL RT-PCR reaction was set up using 5 µL Takara SYBR Premix Ex *Taq* II (Takara Bio Inc.), 0.2 µM each of the gene-specific primer (*CgPDS* qRT F and *CgPDS* qRT R) and 1 µL of first-strand cDNA. The qRT-

PCR was performed using BIO-RAD CFX96 machine. PCR specificity was checked by melting curve analysis, and data were analysed using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

4.2.9 Primers used in this study

Primer Name	Sequence (5′ -> 3′)
*CgPDS*_ALSV_*Xho_F*	TACATCTCGAGTTTGGGGCTTATCCCAAT
*CgPDS*_ALSV_*Bam_R*	TACATGGATCCTCTCATCCATTCTTGCTC
*CgACS*_ALSV_*Xho_F*	TACATCTCGAGCCTTACTATCCTGGATTTGACAGAG
*CgACS*_ALSV_*Bam_R*	TACATGGATCCAATGTCTTCGATTGTGGACCGTTG
R2ALS1363_F	GCGAGGCACTCCTTA
R2ALS1551_R	GCAAGGTGGTCGTGA
CgPDS qRT F	TCAGGAGAAGCATGGCTCTAAG
CgPDS qRT R	TCACCACCCAATGACTGAATATAC
CmACTIN2 F	ATTCTTGCATCTCTAAGTACCTTCC
CmACTIN2 R	CCAACTAAAGGGAAATAACTCACC

4.3 Results

4.3.1 In vitro regeneration and transformation attempts for C. grandis

We attempted to establish *in vitro* regeneration and transformation system for *C. grandis*. Young leaf and tendril explants of *C. grandis* when cultured on callus-induction media resulted in formation of profuse callus after 3 to 4 weeks of incubation (Figure 4-2B, C). Increased callus mass in selection media indicated successful infection and transformation of the explants with *Agrobacterium* GV2260 carrying pBI121 plasmid that confers kanamycin resistance.
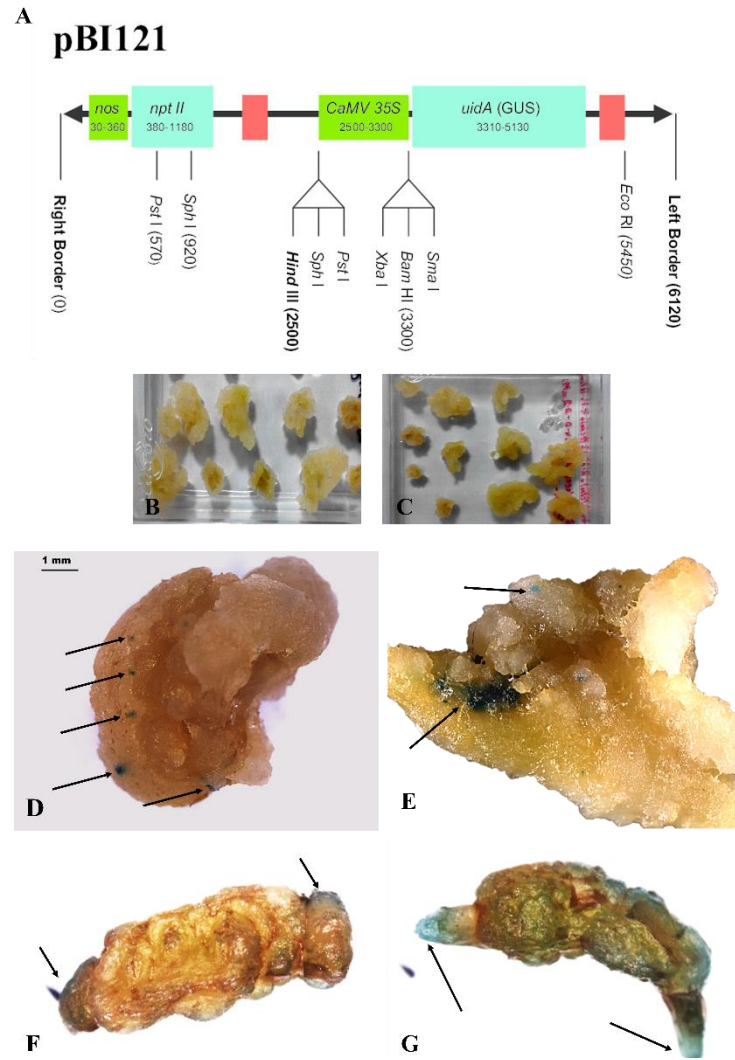
Figure 4-2. *In vitro* **regeneration and transformation attempts with *C. grandis*.** (A) Map for pBI121 binary vector with 35S: GUS maker used for *Agrobacterium*-mediated transformation of *C. grandis*. (B) Calli generated from *C. grandis* leaf explants. (C) Calli generated from *C. grandis* tendrils explants. (D, E) Transgenic calli originated from leaf showing GUS expression. (F, G) Transgenic calli originated from tendril showing GUS expression.

This was confirmed by staining the calli for GUS activity using 5-bromo- 4-chloro-3-indolyl β-D-glucuronide as per Jefferson *et al.* (1987) protocol. Numerous blue spots were observed upon GUS staining of calli originating from both the leaf and tendril explants (Figure 4-2D-G). This indicated multiple sites of *Agrobacterium* infection on each explant. However, these transformed calli did not respond to the shoot induction media tested in this study. We failed to regenerate plantlets from transformed calli even after multiple attempts with different

phytohormone compositions. Hence, we shifted to virus induced gene silencing approach for gene function analysis in *C. grandis*.

4.3.2 Silencing of Phytoene desaturase (PDS) in Nicotiana benthamiana and Coccinia grandis plants using TRV vectors

At the beginning, TRV-based VIGS vectors were confirmed to be functional in *N. benthamiana*. pTRV2-*NtPDS* treated *N. benthamiana* plants showed photobleaching phenotype after around 10 days post-infiltration (Figure 4-3A, B). Also, VIGS was carried out using pTRV2-*CgPDS A* and pTRV2- *CgPDS B* in a total of 29 *C. grandis* plants, of which 10 plants showed yellowing or white patches on leaves to varying degrees (Figure 4-3C). However, none of the *C. grandis* plants showed the typical photobleaching phenotype usually observed upon *PDS* silencing.

4.3.3 Silencing of Pistillata – a B class floral Organ Identity gene using TRV

To achieve silencing of *C. grandis PISTILLATA* (*CgPI*) gene, three fragments of *CgPI* namely; *CgPI A* (219 bp), *CgPI B* (251 bp) and *CgPI C* (630 bp) were amplified from cDNA of wild-type flower and cloned into pTRV2 vector (Figure 4-3A). Fragment A and B sequences were unique to the *CgPI,* whereas fragment C included all the conserved regions. The fragments were designed in such a way that to see if there are any differences in phenotype when both conserved and unique regions are targeted. A total of 20 plants were infiltrated using these three fragments of *CgPI* cloned into pTRV2. In one of the mature plants (3 months old) agroinfiltrated with pTRV2-*CgPI C* construct, flowers with aberrant stamens were observed (Figure 4-3D). The phenotype looked like a chimeric silencing, and some part of the anther had no pollen development (Figure 4-3D). However, we were unable to confirm the silencing of *CgPI* at transcript level, and the phenotype was not reproducible.
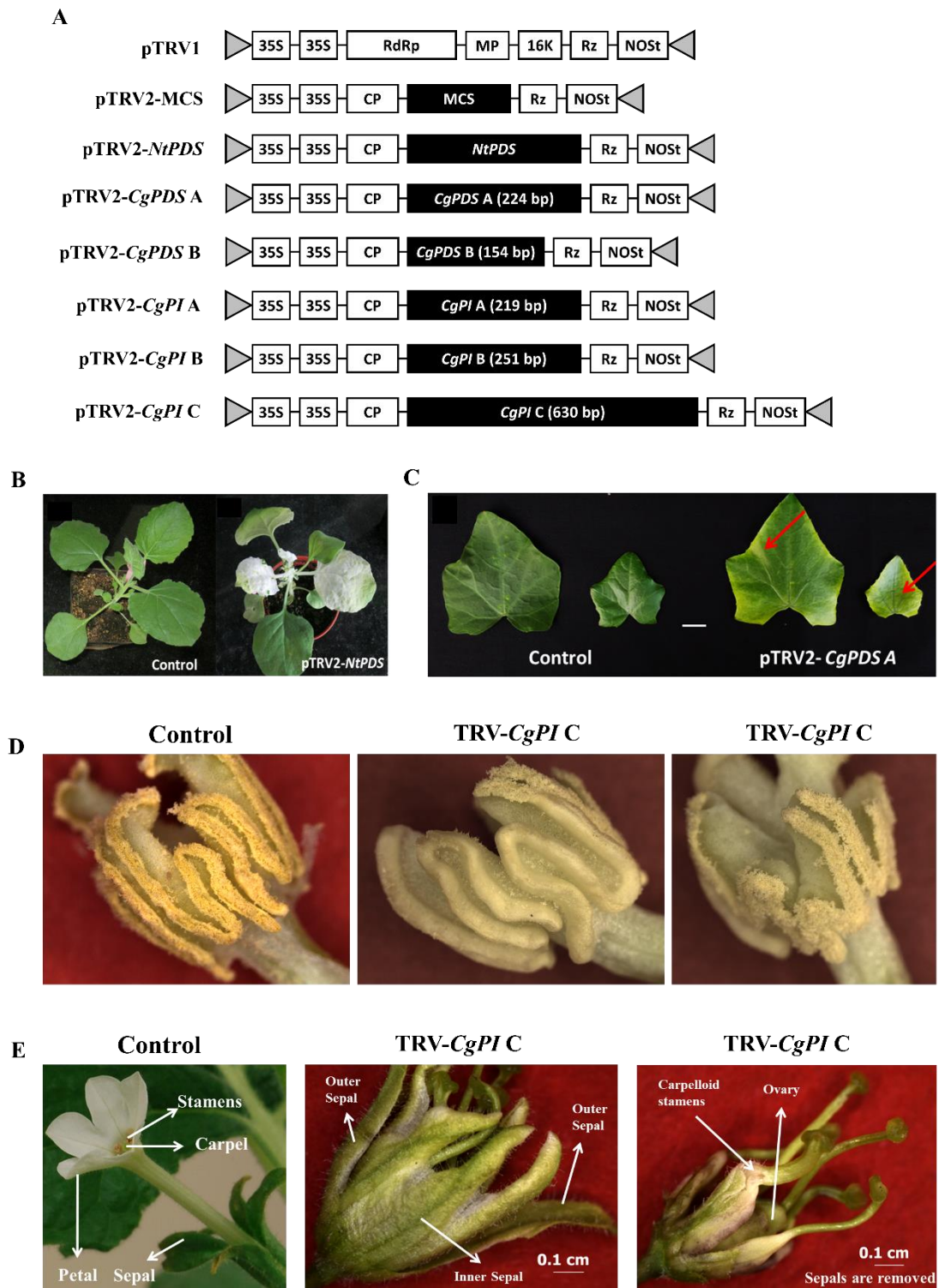
Figure 4-3. **VIGS trials using TRV.** (A) TRV VIGS vectors used in current study. (B) *PDS* silencing in *N. benthamiana* plants. (C) *PDS* silencing in *C. grandis* plants. (D) *CgPI* silencing in *C. grandis* plants. (E) *PI* silencing in *N. benthamiana* plants. Devani and Kushwaha, Unpublished data, Molecular Plant Biology Lab, IISER Pune.

4.3.4 Silencing of Phytoene desaturase in C. grandis using ALSV vectors

Since the TRV-based silencing was not efficient in *C. grandis*, we explored the ALSV approach. To test the efficiency of ALSV vectors in silencing the endogenous genes of *C. grandis*, we again targeted the *phytoene desaturase gene* (*PDS*). A 300 bp fragment of the *CgPDS* gene was cloned into the ALSV-RNA2 vector (Figure 4-4) and the resulting virus (ALSV-*CgPDS*) was mechanically inoculated to *C. quinoa* plants at 6 leaves stage.
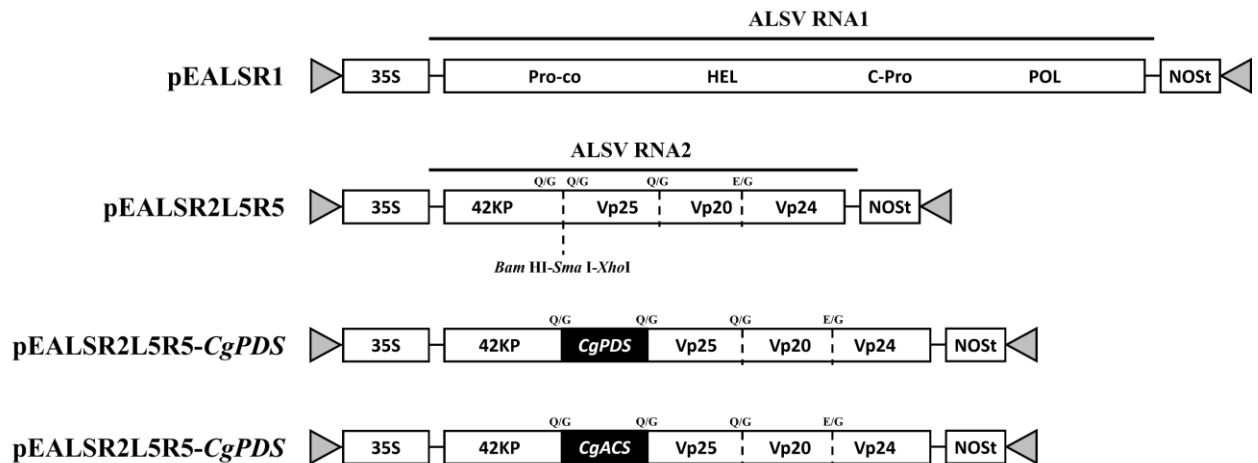
Figure 4-4. **ALSV in pUC-based expression cassette vectors used for VIGS in *C. grandis*.** *CgPDS* and *CgACS* fragments were cloned in the multiple cloning site of PEALSR2L5R5 vector such that they do not disrupt the coding frame of ALSV RNA2.

The leaf titre of the infected *C. quinoa* plants was used as inoculum for treating the *C. grandis* plants (at cotyledon stage). Out of the 20 inoculated *C. grandis* plants, 8 showed the photobleached phenotype, characteristic of *PDS* gene silencing (Figure 4-5A-C, G). The infected plants started to show photobleaching along the veins of the upper un-inoculated leaves, 17-22 days post-inoculation. A uniform photobleaching was observed in the newly developed leaves, petioles and stems of infected *C. grandis* plants indicating that the *PDS* gene was silenced by the ALSV vector efficiently. Silencing of the *PDS* gene persisted for 45-50 days post-inoculation, after which recovery of the silencing phenotype was observed in the newly developing leaves.

In order to confirm that the observed photobleached phenotype was due to the silencing of the *CgPDS* gene by the ALSV vectors, we studied the *CgPDS* expression levels in infected and wild-type *C. grandis* plants. Quantitative RT-PCR (qRT-PCR) was also used to compare the endogenous expression levels of the *PDS* gene between ALSV-*CgPDS*-treated and untreated *C.*

grandis plants (Figure 4-6A). We have also detected the presence of ALSV RNA in the infected plants by RT-PCR. (Figure 4-6B). The presence of ALSV RNA was detected in all the plants that showed photobleached phenotype and in none of the untreated plants.
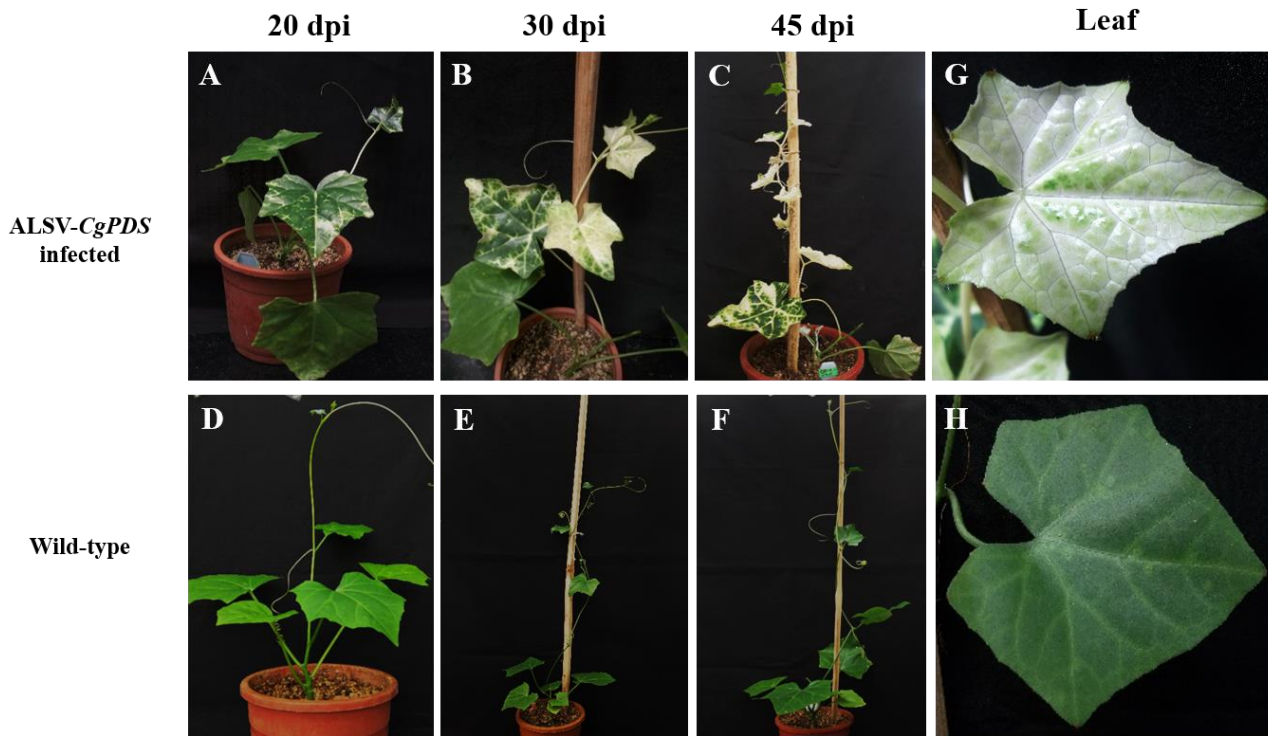
Figure 4-5. **VIGS of *PDS* gene in *Coccinia grandis* by ALSV vector infection.** (A, B, C) *C. grandis* plants infected with pEALSR1+pEALSR2L5R5-*CgPDS* showing photobleached phenotype 20, 30 and 45 days post infection respectively. (D, E, F) Wild-type *C. grandis* plants.

The qRT-PCR results supported our conclusion that the observed photobleached phenotype was due to the reduction in the endogenous levels of *CgPDS* transcripts. The relative levels of the *CgPDS* gene expression was lower in the photobleached plants compared to the untreated plants. We detected an average of 6-fold decrease in the level of *CgPDS* in the ALSV-*CgPDS* treated plants compared to the untreated plants indicating an efficient knock-down of *CgPDS* gene in *C. grandis*.

A

B
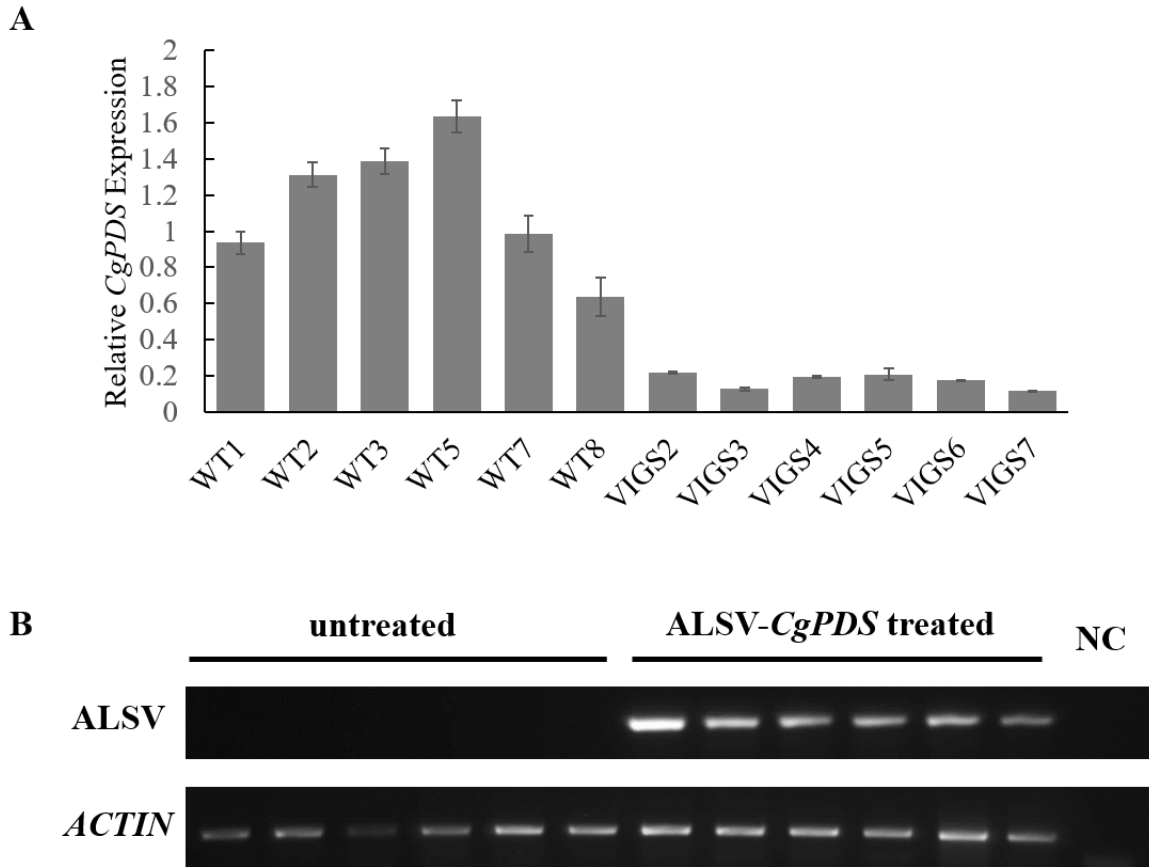
Figure 4-6. **Molecular confirmation of ALSV VIGS-based silencing of *CgPDS* in *C. grandis*.** (A) Relative expression levels of *CgPDS* transcripts in ALSV-*CgPDS* infected and wild-type control plants of *C. grandis*. (B) Detection of ALSV viral RNA in selected ALSV-CgPDS infected and wild-type control plants of *C. grandis* by RT-PCR. Lane NC represents the no template control.

4.3.5 Infection of C. grandis with agroinfiltration compatible ALSV VIGS vectors

Agroinfiltration compatible vectors for ALSV VIGS were developed in our lab to simplify the process of viral infection in *C. grandis* as well as to improve the efficiency of infection. Since the agroinfiltration compatible vectors do not require the step of virus propagation in *C. quinoa*, the total time duration to achieve VIGS in *C. grandis* will also be greatly reduced. Based on the cloning strategy adapted by Kon and Yoshikawa (2014), we developed agroinfiltration compatible ALSV VIGS vectors (pC1302ALSR1 and pC1301ALSR2) in our lab (Figure 4-7A). These vectors were transformed into *Agrobacterium tumefaciens* strain GV2260 separately. Ten seedlings of *C. grandis* at cotyledonary stage were infiltrated with *Agrobacterium* cultures harbouring pC1302ALSR1 & pC1301ALSR2 in 1:1 ratio. RT-PCR was carried out to amplify viral mRNA from the distant leaves of infected plants. We observed

successful infection of virus in 2 out of 10 plants indicating the suitability of agroinfiltration compatible ALSV vectors for inducing VIGS in *C. grandis* (Figure 4-7B).
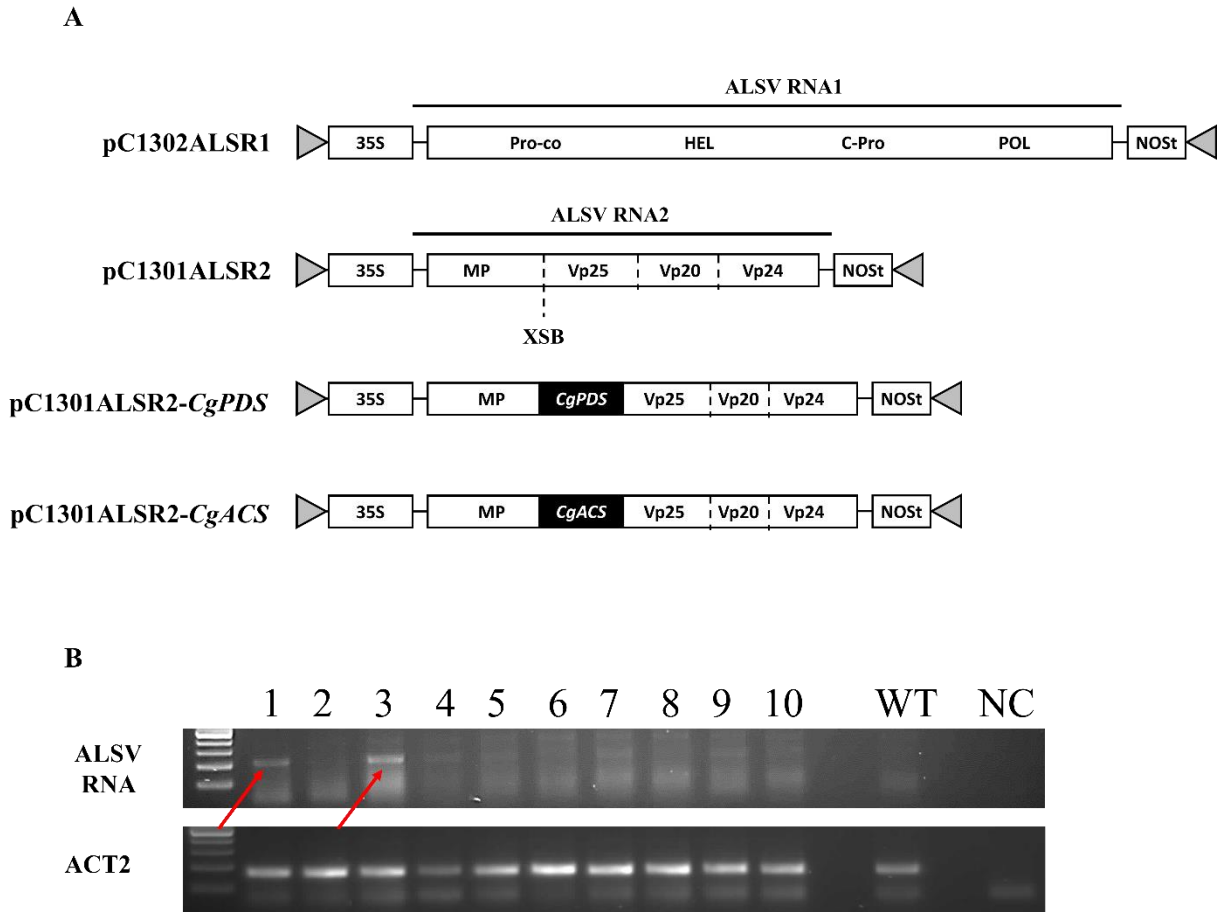
Figure 4-7. **Agroinfiltration compatible ALSV VIGS vectors can infect *C. grandis*.** (A) Binary vectors with ALSV RNA in pCAMBIA cassette developed in our lab using cloning strategy described by Kon & Yoshikawa (2014). *CgPDS* and *CgACS* fragments have also been cloned in pC1301ALSR2 for future VIGS studies. (B) Detection of ALSV viral RNA by RT-PCR from *C. grandis* plants agroinfiltrated with pC1302ALSR1+pC1301ALSR2 in 1:1 ratio.

4.4 Discussion

All our efforts with regard to establishment of an *in vitro* regeneration and transformation system of *C. grandis* remain unsuccessful. Hence, we chose VIGS as an alternative method of gene functional analysis.

4.4.1 Choice of TRV and ALSV as viral vectors for inducing VIGS in C. grandis

The choice of viral vector for VIGS depends on the infectivity of the virus in the plant of interest. Also, many of the candidate genes potentially governing sex expression and differentiation are likely to be expressed in floral meristem or immediately after the floral organ initiation. Most of the viruses are found to be excluded from the actively growing regions of plants such as floral meristem (Burch-Smith *et al.*, 2004). Keeping both the points under consideration, TRV- and ALSV-based VIGS vectors were chosen for studying the candidate genes related to floral development in *C. grandis.* Both TRV and ALSV have been reported to be able to penetrate the actively growing regions of the plant (Ratcliff *et al.*, 2001, Igarashi *et al.*, 2009). Additionally, both TRV and ALSV are known to have a broad host range and hence, we speculated that these will have a better chances of infecting *C. grandis* plants, in which none of the VIGS vectors have so far been tested. ALSV is the only virus reported for virus induced gene silencing in other plants of Cucurbitaceae such as melon cucumber, etc. (Igarashi *et al.*, 2009). Hence, the possibility of ALSV being able to infect and induce VIGS in *C. grandis* might be potentially high*,* which also belongs to Cucurbitaceae. However, one of the disadvantages known to be associated with ALSV vector is the gene expression strategy of the virus genome (Igarashi *et al.*, 2009). As the proteins encoded by ALSV genome were expressed by polyprotein synthesis followed by proteolytic processing, it is necessary to ligate target sequences in frame to the cloning sites of the ALSV vector. Though TRV based vectors were able to infect *C. grandis* plants, the infection efficiency was very low compared to that of ALSV. Also, the strength of *CgPDS* gene knockdown achieved using TRV based vectors was not comparable to that of ALSV based vectors. Hence, we have chosen ALSV as the virus of choice for studying the functions of candidate genes in sex determination and differentiation using VIGS in *C. grandis*.

4.4.2 Efficiency of TRV based vectors for VIGS in C. grandis

Initially, VIGS trials were made with TRV-based vectors in *C. grandis* for silencing *CgPDS* gene. Simultaneously, VIGS attempts were made in *Nicotiana benthamiana* as a positive control using pTRV2-*NtPDS* vector. Silencing of the *PDS* gene leads to decrease in the production of photoprotective carotenoid proteins which in turn results in the breakdown of chlorophyll pigments. Robust photobleaching was observed in the leaves of agroinfiltrated *N. benthamiana* plants (Figure 4-3B). These finding are consistent with previous report of *PDS*

gene silencing in *N. benthamiana* (Kumagai *et al.*, 1995). Two different pTRV2 vectors carrying different regions of *CgPDS* cDNA were used to silence *CgPDS* gene in *C. grandis*. We observed that agroinfiltration of both the vectors resulted in poor infection and silencing. Only minor yellowing and white patches were observed on leaves of some of the agroinfiltrated *C. grandis* plants indicating a poor silencing of *CgPDS* gene (Figure 4-3C). Hence, these results eliminated the effect of nucleotide sequence and/or the position of insert on target gene cloned into pTRV2 vector. Similar experiment was carried out to silence the *CgPI*, a B-class floral organ identity gene in *C. grandis*. Three different fragments of *CgPI* were cloned into pTRV2 vectors separately. Agroinfiltration was carried out in 10 seedlings each of *C. grandis* as well as *N. benthamiana,* where later was used as a positive control. Homeotic transformation expected upon *PI* silencing was observed in the flowers of *N. benthamiana* plants infiltrated with pTRV1+pTRV2-*CgPI* C constructs (Figure 4-3E). The transformation of petals into sepal-like structures as well as transformation of stamen into carpelloid organs was noticed (Figure 4-3C). Since *PI* gene is well conserved among flowering plants, TRV vectors with even the *C. grandis PI* gene fragment was able to silence the *PI* gene in *N. benthamiana* plants. This clearly indicated that the pTRV2 vector constructed with *CgPI* fragment was functional for inducing VIGS in *N. benthamiana*. However, when the *C. grandis* plants were agroinfiltrated with the same vectors, there was no apparent change in flower bud phenotype. Only one *C. grandis* plant had aberrant stamens, where some part of the anthers had no pollen development (Figure 4-3D). This observation was consistent with the results in *Arabidopsis* (a hermaphrodite), where *PISTILLATA* knockdown caused stamens to retain their identity, but pollen development was completely hindered (Wuest *et al.*, 2012). However, we were unable to confirm the silencing at transcript level, and the phenotype was also not reproducible.

4.4.3 Efficiency of ALSV based vectors for VIGS in C. grandis

Unlike TRV, we observed that VIGS using the ALSV vectors harbouring the *CgPDS* gene fragment resulted in robust inhibition of carotenoid biosynthesis in *C. grandis*. This resulted in photo-bleaching phenotypes approximately 15 to 20 days post-inoculation (Figure 4-5A-C, G). Target gene knock-down was confirmed by analysing the expression levels of the *CgPDS* gene in infected *C. grandis* plants (Figure 4-6A). The recovery of the photo-bleached phenotype in the infected plants started after 45-50 days post-inoculation, which can be

attributed to the transient nature of VIGS (Baulcombe, 1999). The control plants treated with ALSV empty vector did not show any apparent symptoms (Figure 4-5D-F, H).

4.4.4 Candidate genes chosen for functional characterization by VIGS

Cucurbitaceae is a large plant family of several species, which have mostly unisexual flowers. Of the 800 species investigated, 460 are monoecious, and 340 are dioecious (Kocyan *et al.*, 2007). In the monoecious members of this family, like *Cucumis melo* (melon) and *Cucumis sativus* (cucumber), the *ACS* (1-aminocyclopropane-1-carboxylic acid synthase) gene is reported to play an important role in sex determination. Expression of the active *ACS* enzyme in female flower buds leads to inhibition of stamen development (Boualem *et al.*, 2008, Boualem *et al.*, 2009). Also, organ identity genes (such as *pistillata* homologs) have implicated in the process of sex determination in plants such as spinach and *Thalictrum* (Di Stilio *et al.*, 2005, Pfent *et al.*, 2005). Hence, we have chosen *CgACS* and *CgPI* as primary candidates for functional characterization with regards to sex expression, using the VIGS protocol that we have standardized in *C. grandis*. If our attempts with these genes turn out to be successful, we plan to expand our VIGS study including few more candidate genes such as *Coccinia* homologs of AMS (ABORTED MICROSPORES), DYT1 (DYSFUNCTIONAL TAPETUM 1), EFE (Ethylene-forming enzyme), etc.

4.4.5 Development of agroinoculation system for ALSV-based VIGS

We were able to achieve successful VIGS in *C. grandis* plants using ALSV in pUC-based expression cassette vector. However, the major limitation of this strategy was that the infectivity of ALSV in pUC-based expression cassette vector was not high enough for direct infection of *C. grandis* plants. Infectious cDNA clones of ALSV were first used for infecting *C. quinoa* plants (highly susceptible to ALSV infection) by mechanical inoculation and the virus propagated in *C. quinoa* was used as an inoculum for infection *C. grandis* as described by (Igarashi *et al.*, 2009). This process is much more complicated compared to the easy induction of RNA silencing using simple agroinfiltration of TRV-based VIGS vectors. Also, the time duration to achieve gene silencing in target plant by mechanical inoculation of ALSV in pUC-based expression cassette vectors is almost 3 times longer compared to that with agroinfiltration of TRV-based VIGS vectors.

To alleviate this issue, an agroinfiltration system has been developed for the ALSV vector by Kon and Yoshikawa (2014). However, these ALSV in pCAMBIA-based binary vectors were not publicly available. Hence, we had to develop similar vectors in our lab and validated the viral propagation in *C. grandis* upon agroinfiltration of these vectors (Figure 4-7). We have also cloned the fragments of *CgPDS* and *CgACS* gene fragments in pCALSR2 base vector (Figure 4-7). Agroinfiltration experiments in *C. grandis* plants are presently ongoing with these vectors to silence the respective genes cloned into them. If we can achieve gene silencing using the binary ALSV vectors, it will greatly aid in the process of functional characterization for candidate genes potentially governing sex determination and differentiation in *C. grandis*.

4.5 Conclusions

In this study, we have reported an efficient protocol for inducing VIGS for the silencing of endogenous genes in *Coccinia grandis* using the ALSV vectors. TRV based VIGS were not found to be efficient in *C. grandis.* The *Agrobacterium*-mediated infiltration method was more efficient in inducing VIGS in the *C. grandis* plants, compared to the mechanical inoculation method using the ALSV vectors. Using the standardized VIGS approach, our future goal is to proceed for the functional characterization of the *CgACS* and *CgPI* gene with respect to their role in sex determination and differentiation in *Coccinia grandis.*

SUMMARY

Investigating sex expression and modification in dioecious *Coccinia grandis* through an integrated transcriptomic and proteomic approach

Sexual systems in plants are almost overwhelmingly diverse. Strikingly, in contrast to animals, most of the species of flowering plants are co-sexual including hermaphrodite species (bearing perfect bisexual flowers) and monoecious species (bearing unisexual male and female flowers on the same individual). However, around 5-6 % of angiosperm species are dioecious (bearing unisexual male and female flowers on separate individuals) (Ming et al., 2011). There are 15600 dioecious angiosperms (5–6% of the total species) in 987 genera (7% of genera) and 175 families (43% of families), with somewhere between 871 to 5000 independent origins of dioecy (Renner, 2014). The patchy distribution of these dioecious species over the phylogenetic tree indicates that dioecy might have evolved multiple times in different families independently (Ainsworth, 2000; Ming et al., 2011). The evolution of dioecy would require occurrence and establishment of at least two independent mutations that include a male-sterility (M → m) mutation and a female-sterility mutation (Su^f → Su^{Female}) (Lewis, 1942; Ross, 1978). Also, these two mutations should be closely linked to prevent the appearance of neuters and hermaphrodites due to their recombination. Eventually, the chromosomes harbouring such mutations will evolve into sex chromosomes. Molecular and genetic mechanisms of sex determination have been studied in many monoecious and dioecious species such as *Silene latifolia* (Hardenack *et al.*, 1994, Lebel-Hardenack *et al.*, 2002, Muyle *et al.*, 2012), *Rumex acetosa* (Kihara and Ono, 1925, Ainsworth *et al.*, 1995, Navajas-Pérez *et al.*, 2005), *Carica papaya* (Liu *et al.*, 2004, Yu *et al.*, 2008, Urasaki *et al.*, 2012), *Asparagus officinalis* (Bracale et al., 1990, Harkess et al., 2015, Tsugama et al., 2017), *Diospyros kaki* (Akagi *et al.*, 2014), *Cucumis sativus* (Boualem *et al.*, 2009, Guo *et al.*, 2010), *Cucumis melo* (Boualem *et al.*, 2008, Martin *et al.*, 2009, Boualem *et al.*, 2015), *Zea mays* (Chuck *et al.*, 2007), *Spinacia oleracea* (Pfent *et al.*, 2005, Sather *et al.*, 2010), *Thalictrum* (Di Stilio *et al.*, 2005) etc. Since it is believed that evolution of dioecy has happened multiple times in angiosperms independent of each other, the molecular mechanisms of achieving unisexual flower development will most probably vary between different groups of dioecious species (Dellaporta and Calderon-Urrea, 1993). Hence, the underlying mechanism of dioecy discovered in one model species may not be conserved in other phylogenetically distant dioecious species.

Coccinia grandis is one such dioecious species having heteromorphic sex chromosomes (Male: 22A+XY & Female: 22A+XX) (Kumar and Viseveshwaraiah, 1952, Bhowmick *et al.*, 2012, Sousa *et al.*, 2013). Y-chromosome of male sex is at least twice the size of the biggest autosome and plays a decisive role in determining maleness (Sousa *et al.*, 2013). *C. grandis* belongs to Cucurbitaceae, a family known for its diverse sexual systems (Kouonon *et al.*, 2009). *Coccinia grandis* has received less attention as a system to understand dioecy compared to *S. latifolia*, *R. acetosa*, *Populus*, etc. Availability of genome sequences for four cucurbit species gives *C. grandis*, an advantage as a system to study dioecy. *C. grandis*, commonly known as ivy gourd, is used as a vegetable and has rich medicinal value.

To understand sex expression and modification in dioecious *C. grandis*, following objectives were selected for the present investigation:

1. Candidate gene approach and comparative transcriptomics to identify novel players involved in sex expression and modification of *C. grandis*.
2. Proteomic characterization of *C. grandis* flower buds in order to understand unisexual flower development and $AgNO_3$ mediated sex modification.
3. Standardization of virus induced gene silencing (VIGS) in *C. grandis* for functional characterization of key differentially expressed genes by VIGS.

Chapter 1: Introduction

In this chapter, we have carried out a thorough literature survey regarding the diversity of sexual forms in plants. The possible routes for evolution of dioecy along with the forces driving the process were summarized. Evolution of sex chromosomes, molecular mechanisms controlling the sex expression in some of the plant systems, reports of chemically induced sex modification were also highlighted in this chapter. We observed that Cucurbitaceae family exhibits considerable variety of sexual systems, wherein many of the species are of agricultural importance. Moreover, availability of genome sequence for 4 different species gives an advantage to Cucurbitaceae for studying the genetics of sex determination. Also, we find that sex determination mechanisms are well-studied in monoecious cucurbits such as melon, cucumber, watermelon, etc. (Zhang *et al.*, 2014c). Recently, a model explaining possible route for evolution of dioecy from monoecy has also been proposed based on the studies in melon by Dr. Bendahmane's group at IPS2, France (Boualem *et al.*, 2015). These authors managed to engineer

dioecious mating system from monoecious melon by fixation of *Cmacs11* loss-of-function alleles at the population level and the maintenance of the dominant allele of *CmWIP1* at the heterozygous level in male plants. However, sex determination mechanisms are not well understood in dioecious members of Cucurbitaceae. Hence, we chose to develop *C. grandis*, a dioecious cucurbit as a model system to understand the molecular mechanisms controlling sex expression.

Since we were interested in establishing *Coccinia grandis* as a model system to understand dioecy, at the beginning, we have carried out detailed morphological, histological and cytological studies with three different sexual forms of *C. grandis* (Ghadge *et al.*, 2014).

Following are the significant findings from this study:

a. A rare gynomonoecious (GyM) form of *C. grandis* bearing pistillate (GyM-F) and morphologically hermaphrodite flowers (GyM-H) was found during our investigation.

b. Morphometric analysis showed the presence of staminodes in flowers of female plant and histological study revealed the absence of carpel initials in the flowers of male plant.

c. Though GyM plant had XX sex chromosomes, the development of stamens occurred in morphologically hermaphrodite flowers (GyM-H) in addition to carpels.

d. Interestingly, the sex of *C. grandis* female plants was modified upon treatment with $AgNO_3$, leading to the development of morphologically hermaphrodite flowers (Ag-H).

e. Interestingly, the pollens of Ag-H and GyM-H flowers were found to be sterile, indicating the probable role of Y-chromosome in governing pollen fertility.

Chapter 2: Candidate gene approach and comparative transcriptomics to identify novel players involved in sex expression and modification of *C. grandis*.

Due to the unavailability of genome sequence and the slow pace at which sex-linked genes are identified, sex expression and modification in *C. grandis* are not well understood. In the beginning, a candidate gene approach was taken to isolate *C. grandis* homologs for sex-determining genes of other species using degenerate primers. Finally, we have carried out a comprehensive RNA-Seq study from early-staged male, female, GyM-H, and Ag-H as well as middle-staged male and GyM-H flower buds (Devani *et al.*, 2017). A *de novo* transcriptome was assembled using Trinity and annotated by BLAST2GO and Trinotate pipelines. The assembled

transcriptome consisted of 467,233 'Trinity Transcripts' clustering into 78,860 'Trinity Genes'. Female_Early_vs_Male_Early, Ag_Early_vs_Female_Early, and GyM-H_Middle_vs_Male_Middle comparisons exhibited 35,694, 3574, and 14,954 differentially expressed transcripts respectively. Further, qRT-PCR analysis of selected candidate genes validated digital gene expression profiling results.

Following are the significant findings from this study:

a. Homologs of candidate genes *PI*, *AG* and *ACS* were isolated from *C. grandis* by degenerate primer approach. Also, miR172 was isolated from *C. grandis* by stem-loop PCR.

b. Expression analysis of these candidate genes revealed male-biased expression of *CgPI*. Interestingly, *CgACS* expression was approx. 10-fold higher in female buds compared to male. This pattern was similar to the expression profiles of *CmACS*-7 (melon) and *CsACS2* (cucumber) indicating a possibility of conserved sex determination mechanisms.

c. Ethylene response-related genes were found to be upregulated in female buds compared to male buds.

d. Also, we observed that $AgNO_3$ treatment suppressed ethylene responses in Ag-H flowers by downregulation of ethylene-responsive transcription factors leading to stamen development.

e. Suppression of ethylene responses in both male and Ag-H compared to female buds suggests a probable role of ethylene in stamen suppression similar to monoecious cucurbits such as melon and cucumber.

f. Further, GO terms related to stamen development were enriched in early-staged male, GyM-H, and Ag-H buds compared to female buds supporting the fact that stamen growth gets arrested in female flowers.

g. Also, pollen fertility associated GO terms were depleted in middle-staged GyM-H buds compared to male buds indicating the necessity of Y-chromosome for pollen fertility.

Chapter 3: Proteomic characterization of *C. grandis* flower buds in order to understand unisexual flower development and AgNO₃ mediated sex modification.

Knowledge of sex-related differences can enhance our understanding about the molecular basis of unisexual flower development. *Coccinia grandis* is a dioecious species belonging to Cucurbitaceae, a family well-known for diverse sexual systems. Male and female plants have 22A+XY and 22A+XX-chromosomes respectively. Previously, we have reported a gynomonoecious form (22A+XX) of *C.* grandis bearing morphologically hermaphrodite flowers (GyM-H). Also, we showed that foliar spray of silver nitrate on female plant induces morphologically hermaphrodite bud development (Ag-H) despite the absence of Y-chromosome. To identify sex-related differences, total proteome from male, female, GyM-H and Ag-H flower buds at early and middle stages of development were analysed by label-free proteomics (Devani *et al.*, 2018: Under review). 3385 proteins were detected (FDR ≤ 1) using *C. grandis* flower bud transcriptome database. SWATH-MS based comparative abundance analysis of 2262 proteins between Female_Early vs Male_Early, Ag_Early vs Female_Early, GyM-H_Middle vs Male_Middle and Ag_Middle vs Male_Middle identified 644, 849, 669 and 591 differentially expressed proteins, respectively.

Following are the significant findings from this study:

a. Upregulation of ethylene biosynthesis-related proteins in female buds compared to male buds indicated a conserved stamen arrest mechanism similar to monoecious cucurbits.

b. AgNO₃ treatment induced proteins related to pollen development in Ag-H buds.

c. Few proteins governing pollen germination and tube growth were highly expressed in male buds compared to Ag-H and GyM-H buds indicating the role of Y-chromosome in governing pollen fertility.

Chapter 4: Standardization of virus induced gene silencing (VIGS) in *C. grandis* for functional characterization of key differentially expressed genes by VIGS.

Our transcriptomic and proteomic studies have identified many interesting sex-biased genes/proteins which might govern the stamen inhibition in female flowers and pollen fertility in male flowers of *C. grandis*. However, lack of an efficient *in vitro* regeneration and genetic transformation tool for *C. grandis* was a major limitation to study the function of the sex-biased

genes identified in our study. Initially, we attempted to achieve *in vitro* regeneration and transformation of *C. grandis* in our lab. *Agrobacterium*-mediated infection was achieved in the leaf and tendril explants of *C. grandi*s, which was confirmed by GUS staining. However, we failed to regenerate shoots from these transformed calli even after testing various growth hormone concentrations. Hence, we chose to standardize virus induced gene silencing tool as an alternative strategy for gene knockdown in *C. grandis* using tobacco rattle virus (TRV) and apple latent spherical virus (ALSV).

Following are the significant findings:

a. TRV based VIGS system was not found to be efficient for gene knock-down in *C. grandis.*

b. We have developed an efficient ALSV-VIGS system for *C. grandis* in order to understand gene function by knock-down (Devani *et al.*, 2018: Under preparation).

c. Robust silencing of *CgPDS* gene in *C. grandis* plants using ALSV-VIGS system was achieved (Devani *et al.*, 2018: Under preparation).

d. Agroinfiltration compatible binary vectors for ALSV VIGS were developed and tested for infection in *C. grandis*.

e. Attempts are in progress to study candidate DE genes from our transcriptomics and proteomics approach using ALSV-VIGS in order to understand their role in sex expression and modification.

Future directions

Overall, our study has enabled identification of new sex-biased genes and provided a basis for functional characterization to understand their role in stamen arrest, pollen fertility, and AgNO$_3$-mediated sex modification. Following future directions can be taken to enhance the knowledge about sex expression and modification in *C. grandis*:

• Analysis of candidate DE genes/proteins identified from transcriptomics and proteomics study using the ALSV-VIGS tool could help us understand their roles in sex determination and differentiation of *C. grandis*.

- *In situ* hybridization studies and gene function analysis for *CgACS* gene would give us an idea regarding the role of ethylene in sex determination in *C. grandis*.

- Detection of *Gyno-h*AT transposon in vicinity of *WIP1* gene. A female specific occurrence of Gyno-*h*AT could shed light on the possible mechanism of stamen arrest in female flowers. Absence of *Gyno-h*AT in male genome can also explain the lack of carpel primordia in male flowers.

- Identification and knockdown of *C. grandis* homolog of *CmACS11* can help us understand the evolution of dioecious cucurbits from monoecious cucurbits.

REFERENCES

References

Ainsworth, C. (2000) Boys and Girls Come Out to Play: The Molecular Biology of Dioecious Plants. *Annals of Botany*, **86**, 211-221.

Ainsworth, C., Crossley, S., Buchanan-Wollaston, V., Thangavelu, M. and Parker, J. (1995) Male and female flowers of the dioecious plant sorrel show different patterns of MADS box gene expression. *The Plant Cell*, **7**, 1583.

Ainsworth, C., Parker, j. and Buchanan-Wollaston, V. (1997a) 5 Sex Determination in Plants. In *Current Topics in Developmental Biology* (Pedersen, R.A. and Schatten, G.P. eds): Academic Press, pp. 167-223.

Ainsworth, C., Parker, j. and Buchanan-Wollaston, V. (1997b) Sex Determination in Plants. In *Current Topics in Developmental Biology* (Pedersen, R.A. and Schatten, G.P. eds): Academic Press, pp. 167-223.

Akagi, T., Henry, I.M., Tao, R. and Comai, L. (2014) A Y-chromosome–encoded small RNA acts as a sex determinant in persimmons. *Science*, **346**, 646.

Al-Kaff, N.S., Covey, S.N., Kreike, M.M., Page, A.M., Pinder, R. and Dale, P.J. (1998) Transcriptional and Posttranscriptional Plant Gene Silencing in Response to a Pathogen. *Science*, **279**, 2113.

Albrecht, C., Russinova, E., Hecht, V., Baaijens, E. and de Vries, S. (2005) The *Arabidopsis thaliana* SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 Control Male Sporogenesis. *The Plant Cell*, **17**, 3337.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403-410.

Andème Ondzighi, C., Christopher, D.A., Cho, E.J., Chang, S.-C. and Staehelin, L.A. (2008) *Arabidopsis* Protein Disulfide Isomerase-5 Inhibits Cysteine Proteases during Trafficking to Vacuoles before Programmed Cell Death of the Endothelium in Developing Seeds. *The Plant Cell*, **20**, 2205.

Baker, H.G. (1959) Reproductive Methods as Factors in Speciation in Flowering Plants. *Cold Spring Harbor Symposia on Quantitative Biology*, **24**, 177-191.

Banks, J.A. (2008) MicroRNA, sex determination and floral meristem determinacy in maize. *Genome Biology*, **9**, 204.

Baulcombe, D.C. (1999) Fast forward genetics based on virus-induced gene silencing. *Current Opinion in Plant Biology*, **2**, 109-113.

Bawa, K.S. (1980) Evolution of Dioecy in Flowering Plants. *Annual Review of Ecology and Systematics*, **11**, 15-39.

Becker, A. and Lange, M. VIGS – genomics goes functional. *Trends in Plant Science*, **15**, 1-4.

Berg, M., Rogers, R., Muralla, R. and Meinke, D. (2005) Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in *Arabidopsis*. *The Plant Journal*, **44**, 866-878.

Bergero, R. and Charlesworth, D. (2011) Preservation of the Y Transcriptome in a 10-Million-Year-Old Plant Sex Chromosome System. *Current Biology*, **21**, 1470-1474.

Berkowitz, O., Jost, R., Pollmann, S. and Masle, J. (2008) Characterization of TCTP, the Translationally Controlled Tumor Protein, from *Arabidopsis thaliana*. *The Plant Cell*, **20**, 3430-3447.

Bernal, A.J., Yoo, C.-M., Mutwil, M., Jensen, J.K., Hou, G., Blaukopf, C., Sørensen, I., Blancaflor, E.B., Scheller, H.V. and Willats, W.G.T. (2008) Functional Analysis of the Cellulose Synthase-Like Genes *CSLD1*, *CSLD2* and *CSLD4* in Tip-Growing *Arabidopsis* Cells. *Plant Physiology*, **148**, 1238.

Bertin, R.I. and Christian, M.N. (1993) Dichogamy in Angiosperms. *Botanical Review*, **59**, 112-152.

Beyer, E.M. (1976) A Potent Inhibitor of Ethylene Action in Plants. *Plant Physiology*, **58**, 268.

Bhaduri, P.N. and Bose, P.C. (1947) Cyto-genetical investigations in some common cucurbits, with special reference to fragmentation of chromosomes as a physical basis of speciation. *Journal of Genetics*, **48**, 237.

Bhowmick, B.K., Jha, T.B. and Jha, S. (2012) Chromosome analysis in the dioecious cucurbit *Coccinia grandis* (L.) Voigt. *Chromosome Science*, **15**, 9-15.

Bhuskute, S., Makde, K. and Deshpande, P. (1986) Staminal organization in *Coccinia grandis* (L.) Voigt. *Annals of botany*, 415-418.

Birke, H., Heeg, C., Wirtz, M. and Hell, R. (2013) Successful fertilization requires the presence of at least one major *O*-acetylserine(thiol)lyase for cysteine synthesis in pollen of *Arabidopsis*. *Plant Physiology*.

Blackburn, K.B. (1923) Sex Chromosomes in Plants. *Nature*, **112**, 687.

Boavida, L.C., Shuai, B., Yu, H.-J., Pagnussat, G.C., Sundaresan, V. and McCormick, S. (2009) A Collection of *Ds* Insertional Mutants Associated With Defects in Male Gametophyte Development and Function in *Arabidopsis thaliana*. *Genetics*, **181**, 1369-1385.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.

Borg, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M. and Twell, D. (2011) The R2R3 MYB Transcription Factor DUO1 Activates a Male Germline-Specific Regulon Essential for Sperm Cell Differentiation in *Arabidopsis*. *The Plant Cell*, **23**, 534.

Boualem, A., Fergany, M., Fernandez, R., Troadec, C., Martin, A., Morin, H., Sari, M.-A., Collin, F., Flowers, J.M., Pitrat, M., Purugganan, M.D., Dogimont, C. and Bendahmane, A. (2008) A Conserved Mutation in an Ethylene Biosynthesis Enzyme Leads to Andromonoecy in Melons. *Science*, **321**, 836.

Boualem, A., Lemhemdi, A., Sari, M.-A., Pignoly, S., Troadec, C., Abou Choucha, F., Solmaz, I., Sari, N., Dogimont, C. and Bendahmane, A. (2016) The Andromonoecious Sex Determination Gene Predates the Separation of *Cucumis* and *Citrullus Genera*. *PLOS ONE*, **11**, e0155444.

Boualem, A., Troadec, C., Camps, C., Lemhemdi, A., Morin, H., Sari, M.-A., Fraenkel-Zagouri, R., Kovalski, I., Dogimont, C., Perl-Treves, R. and Bendahmane, A. (2015) A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. *Science*, **350**, 688.

Boualem, A., Troadec, C., Kovalski, I., Sari, M.-A., Perl-Treves, R. and Bendahmane, A. (2009) A Conserved Ethylene Biosynthesis Enzyme Leads to Andromonoecy in Two *Cucumis* Species. *PLOS ONE*, **4**, e6144.

Boyce, R., Chilana, P. and Rose, T.M. (2009) iCODEHOP: a new interactive program for designing COnsensus-DEgenerate Hybrid Oligonucleotide Primers from multiply aligned protein sequences. *Nucleic Acids Research*, **37**, W222-W228.

Bracale, M., Galli, M., Falavigna, A. and Soave, C. (1990) Sexual differentiation in *Asparagus officinalis* L. *Sexual Plant Reproduction*, **3**, 23-30.

Brigneti, G., Martín-Hernández, A.M., Jin, H., Chen, J., Baulcombe, D.C., Baker, B. and Jones, J.D.G. (2004) Virus-induced gene silencing in *Solanum* species. *The Plant Journal*, **39**, 264-272.

Bruun-Rasmussen, M., Madsen, C.T., Jessing, S. and Albrechtsen, M. (2007) Stability of Barley stripe mosaic virus–Induced Gene Silencing in Barley. *Molecular Plant-Microbe Interactions*, **20**, 1323-1331.

Burch-Smith, T.M., Anderson, J.C., Martin, G.B. and Dinesh-Kumar, S.P. (2004) Applications and advantages of virus-induced gene silencing for gene function studies in plants. *The Plant Journal*, **39**, 734-746.

Cai, X.-Z., Xu, Q.-F., Wang, C.-C. and Zheng, Z. (2006) Development of a Virus-Induced Gene-Silencing System for Functional Analysis of the RPS2-Dependent Resistance Signalling Pathways in *Arabidopsis*. *Plant Molecular Biology*, **62**, 223-232.

Cakir, C. and Tör, M. (2010) Factors influencing Barley Stripe Mosaic Virus-mediated gene silencing in wheat. *Physiological and Molecular Plant Pathology*, **74**, 246-253.

Causier, B., Schwarz-Sommer, Z. and Davies, B. (2010) Floral organ identity: 20 years of ABCs. *Seminars in Cell & Developmental Biology*, **21**, 73-79.

Chakravorti, A. (1948) Cytology of *Coccinia indica* W. & A. with reference to the behaviour of its sex-chromosomes. In *Proceedings of the Indian Academy of Sciences-Section B*: Springer, pp. 74-86.

Chang, Z., Li, G., Liu, J., Zhang, Y., Ashby, C., Liu, D., Cramer, C.L. and Huang, X. (2015) Bridger: a new framework for de novo transcriptome assembly using RNA-seq data. *Genome Biology*, **16**, 30.

Chapman, S., Kavanagh, T. and Baulcombe, D. (1992) Potato virus X as a vector for gene expression in plants. *The Plant Journal*, **2**, 549-557.

Charlesworth, B. and Charlesworth, D. (1978) A Model for the Evolution of Dioecy and Gynodioecy. *The American Naturalist*, **112**, 975-997.

Charlesworth, D. (2002) Plant sex determination and sex chromosomes. *Heredity*, **88**, 94.

Charlesworth, D. (2013) Plant sex chromosome evolution. *Journal of Experimental Botany*, **64**, 405-420.

Charlesworth, D. and Guttman, D.S. (1999) The evolution of dioecy and plant sex chromosome systems. *Sex determination in plants*, 25-49.

Charnov, E.L. and Bull, J. (1977) When is sex environmentally determined? *Nature*, **266**, 828.

Chen, H., Sun, J., Li, S., Cui, Q., Zhang, H., Xin, F., Wang, H., Lin, T., Gao, D., Wang, S., Li, X., Wang, D., Zhang, Z., Xu, Z. and Huang, S. (2016) An *ACC Oxidase* Gene Essential for Cucumber Carpel Development. *Molecular Plant*, **9**, 1315-1327.

Chibalina, Margarita V. and Filatov, Dmitry A. (2011) Plant Y Chromosome Degeneration Is Retarded by Haploid Purifying Selection. *Current Biology*, **21**, 1475-1479.

Chuck, G., Meeley, R., Irish, E., Sakai, H. and Hake, S. (2007) The maize *tasselseed4* microRNA controls sex determination and meristem cell fate by targeting *Tasselseed6/indeterminate spikelet1*. *Nature Genetics*, **39**, 1517.

Covey, S.N., Al-Kaff, N.S., Lángara, A. and Turner, D.S. (1997) Plants combat infection by gene silencing. *Nature*, **385**, 781.

Darwin, C. (1876) *The effects of cross and self fertilisation in the vegetable kingdom*: J. Murray.

de Azevedo Souza, C., Kim, S.S., Koch, S., Kienow, L., Schneider, K., McKim, S.M., Haughn, G.W., Kombrink, E. and Douglas, C.J. (2009) A Novel Fatty Acyl-CoA Synthetase Is Required for Pollen Development and Sporopollenin Biosynthesis in *Arabidopsis*. *The Plant Cell*, **21**, 507-525.

De Storme, N. and Geelen, D. (2011) The Arabidopsis Mutant *jason* Produces Unreduced First Division Restitution Male Gametes through a Parallel/Fused Spindle Mechanism in Meiosis II. *Plant Physiology*, **155**, 1403.

Delichère, C., Veuskens, J., Hernould, M., Barbacar, N., Mouras, A., Negrutiu, I. and Monéger, F. (1999) *SlY1*, the first active gene cloned from a plant Y chromosome, encodes a WD-repeat protein. *The EMBO Journal*, **18**, 4169.

Dellaporta, S.L. and Calderon-Urrea, A. (1993) Sex determination in flowering plants. *The Plant Cell*, **5**, 1241.

Delmas, F., Séveno, M., Northey, J.G.B., Hernould, M., Lerouge, P., McCourt, P. and Chevalier, C. (2008) The synthesis of the rhamnogalacturonan II component 3-deoxy-D-manno-2-octulosonic acid (Kdo) is required for pollen tube growth and elongation. *Journal of Experimental Botany*, **59**, 2639-2647.

Dettmer, J., Schubert, D., Calvo-Weimar, O., Stierhof, Y.-D., Schmidt, R. and Schumacher, K. (2005) Essential role of the V-ATPase in male gametophyte development. *The Plant Journal*, **41**, 117-124.

Devani, R.S., Sinha, S., Banerjee, J., Sinha, R.K., Bendahmane, A. and Banerjee, A.K. (2017) *De novo* transcriptome assembly from flower buds of dioecious, gynomonoecious and chemically masculinized female *Coccinia grandis* reveals genes associated with sex expression and modification. *BMC Plant Biology*, **17**, 241.

Di Stilio, V.S., Kramer, E.M. and Baum, D.A. (2005) Floral MADS box genes and homeotic gender dimorphism in *Thalictrum dioicum* (Ranunculaceae) – a new model for the study of dioecy. *The Plant Journal*, **41**, 755-766.

Ding, S.-W. and Voinnet, O. (2007) Antiviral Immunity Directed by Small RNAs. *Cell*, **130**, 413-426.

Dobritsa, A.A., Geanconteri, A., Shrestha, J., Carlson, A., Kooyers, N., Coerper, D., Urbanczyk-Wochniak, E., Bench, B.J., Sumner, L.W., Swanson, R. and Preuss, D. (2011) A Large-Scale Genetic Screen in *Arabidopsis* to Identify Genes Involved in Pollen Exine Production. *Plant Physiology*, **157**, 947-970.

Dobritsa, A.A., Lei, Z., Nishikawa, S.-i., Urbanczyk-Wochniak, E., Huhman, D.V., Preuss, D. and Sumner, L.W. (2010) *LAP5* and *LAP6* Encode Anther-Specific Proteins with Similarity to Chalcone Synthase Essential for Pollen Exine Development in *Arabidopsis*. *Plant Physiology*, **153**, 937-955.

Donaire, L., Barajas, D., Martínez-García, B., Martínez-Priego, L., Pagán, I. and Llave, C. (2008) Structural and Genetic Requirements for the Biogenesis of Tobacco Rattle Virus-Derived Small Interfering RNAs. *Journal of Virology*, **82**, 5167-5177.

Dong, X., Hong, Z., Sivaramakrishnan, M., Mahfouz, M. and Verma, D.P.S. (2005) Callose synthase (CalS5) is required for exine formation during microgametogenesis and for pollen viability in *Arabidopsis*. *The Plant Journal*, **42**, 315-328.

Dong, Y., Burch-Smith, T.M., Liu, Y., Mamillapalli, P. and Dinesh-Kumar, S.P. (2007) A Ligation-Independent Cloning Tobacco Rattle Virus Vector for High-Throughput Virus-Induced Gene Silencing Identifies Roles for *NbMADS4-1* and *2* in Floral Development. *Plant Physiology*, **145**, 1161.

Farbos, I., Veuskens, J., Vyskot, B., Oliveira, M., Hinnisdaels, S., Aghmir, A., Mouras, A. and Negrutiu, I. (1999) Sexual Dimorphism in White Campion: Deletion on the Y Chromosome Results in a Floral Asexual Phenotype. *Genetics*, **151**, 1187.

Filatov, D.A., Monéger, F., Negrutiu, I. and Charlesworth, D. (2000) Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution. *Nature*, **404**, 388.

Forsthoefel, N.R. and Vernon, D.M. (2011) Effect of sporophytic *PIRL9* genotype on post-meiotic expression of the *Arabidopsis pirl1*;*pirl9* mutant pollen phenotype. *Planta*, **233**, 423-431.

Fredslund, J., Schauser, L., Madsen, L.H., Sandal, N. and Stougaard, J. (2005) PriFi: using a multiple alignment of related sequences to find primers for amplification of homologs. *Nucleic Acids Research*, **33**, W516-W520.

Freeman, D.C., Doust, J.L., El-Keblawy, A., Miglia, K.J. and McArthur, E.D. (1997) Sexual specialization and inbreeding avoidance in the evolution of dioecy. *The Botanical Review*, **63**, 65-92.

Fu, D.Q., Zhu, B.Z., Zhu, H.L., Jiang, W.B. and Luo, Y.B. (2005) Virus-induced gene silencing in tomato fruit. *The Plant Journal*, **43**, 299-308.

Ghadge, A.G., Karmakar, K., Devani, R.S., Banerjee, J., Mohanasundaram, B., Sinha, R.K., Sinha, S. and Banerjee, A.K. (2014) Flower development, pollen fertility and sex

expression analyses of three sexual phenotypes of *Coccinia grandis*. *BMC Plant Biology*, **14**, 325.

Golan-Goldhirsh, A., Peri, I., Birk, Y. and Smirnoff, P. (1998) Inflorescence bud proteins of *Pistacia vera*. *Trees*, **12**, 415-419.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J. and Conesa, A. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, **36**, 3420-3435.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. and Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644.

Grienenberger, E., Kim, S.S., Lallemand, B., Geoffroy, P., Heintz, D., Souza, C.d.A., Heitz, T., Douglas, C.J. and Legrand, M. (2010) Analysis of *TETRAKETIDE α-PYRONE REDUCTASE* Function in *Arabidopsis thaliana* Reveals a Previously Unknown, but Conserved, Biochemical Pathway in Sporopollenin Monomer Biosynthesis. *The Plant Cell*, **22**, 4067-4083.

Gu, J.N., Zhu, J., Yu, Y., Teng, X.D., Lou, Y., Xu, X.F., Liu, J.L. and Yang, Z.N. (2014) DYT1 directly regulates the expression of *TDF1* for tapetum development and pollen wall formation in *Arabidopsis*. *The Plant Journal*, **80**, 1005-1013.

Guo, S., Zheng, Y., Joung, J.-G., Liu, S., Zhang, Z., Crasta, O.R., Sobral, B.W., Xu, Y., Huang, S. and Fei, Z. (2010) Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. *BMC Genomics*, **11**, 384.

Guttman, D.S. and Charlesworth, D. (1998) An X-linked gene with a degenerate Y-linked homologue in a dioecious plant. *Nature*, **393**, 263.

Gygi, S.P., Rochon, Y., Franza, B.R. and Aebersold, R. (1999) Correlation between Protein and mRNA Abundance in Yeast. *Molecular and Cellular Biology*, **19**, 1720-1730.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N. and Regev, A. (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, **8**, 1494.

Hardenack, S., Ye, D., Saedler, H. and Grant, S. (1994) Comparison of *MADS box* gene expression in developing male and female flowers of the dioecious plant white campion. *The Plant Cell*, **6**, 1775.

Harkess, A., Mercati, F., Shan, H.-Y., Sunseri, F., Falavigna, A. and Leebens-Mack, J. (2015) Sex-biased gene expression in dioecious garden asparagus (*Asparagus officinalis*). *New Phytologist*, **207**, 883-892.

Igarashi, A., Yamagata, K., Sugai, T., Takahashi, Y., Sugawara, E., Tamura, A., Yaegashi, H., Yamagishi, N., Takahashi, T., Isogai, M., Takahashi, H. and Yoshikawa, N. (2009) Apple latent spherical virus vectors for reliable and effective virus-induced gene silencing among a broad range of plants including tobacco, tomato, *Arabidopsis thaliana*, cucurbits, and legumes. *Virology*, **386**, 407-416.

Isaacson, T., Damasceno, C.M.B., Saravanan, R.S., He, Y., Catalá, C., Saladié, M. and Rose, J.K.C. (2006) Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nature Protocols*, **1**, 769.

Ito, T., Nagata, N., Yoshiba, Y., Ohme-Takagi, M., Ma, H. and Shinozaki, K. (2007) *Arabidopsis MALE STERILITY1* Encodes a PHD-Type Transcription Factor and Regulates Pollen and Tapetum Development. *The Plant Cell*, **19**, 3549.

Ito, T. and Shinozaki, K. (2002) The *MALE STERILITY1* Gene of *Arabidopsis*, Encoding a Nuclear Protein with a PHD-finger Motif, is Expressed in Tapetal Cells and is Required for Pollen Maturation. *Plant and Cell Physiology*, **43**, 1285-1292.

Janoušek, B., Široký, J. and Vyskot, B. (1996) Epigenetic control of sexual phenotype in a dioecious plant,Melandrium album. *Molecular and General Genetics MGG*, **250**, 483-490.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal*, **6**, 3901-3907.

Jia, G., Liu, X., Owen, H.A. and Zhao, D. (2008) Signaling of cell fate determination by the TPD1 small protein and EMS1 receptor kinase. *Proceedings of the National Academy of Sciences*, **105**, 2220.

Jiang, L., Yang, S.-L., Xie, L.-F., Puah, C.S., Zhang, X.-Q., Yang, W.-C., Sundaresan, V. and Ye, D. (2005) *VANGUARD1* Encodes a Pectin Methylesterase That Enhances Pollen Tube Growth in the *Arabidopsis* Style and Transmitting Tract. *The Plant Cell*, **17**, 584.

Johnson, M.A., von Besser, K., Zhou, Q., Smith, E., Aux, G., Patton, D., Levin, J.Z. and Preuss, D. (2004) Arabidopsis *hapless* Mutations Define Essential Gametophytic Functions. *Genetics*, **168**, 971-982.

Kalantidis, K., Schumacher, H.T., Alexiadis, T. and Helm, J.M. (2008) RNA silencing movement in plants. *Biology of the Cell*, **100**, 13-26.

Kihara, H. and Ono, T. (1925) The sex-chromosomes of *Rumex acetosa*. *Zeitschrift für induktive Abstammungs-und Vererbungslehre*, **39**, 1-7.

Kim, C.Y., Bove, J. and Assmann, S.M. (2008) Overexpression of wound-responsive RNA-binding proteins induces leaf senescence and hypersensitive-like cell death. *New Phytologist*, **180**, 57-70.

Kim, S.S., Grienenberger, E., Lallemand, B., Colpitts, C.C., Kim, S.Y., Souza, C.d.A., Geoffroy, P., Heintz, D., Krahn, D., Kaiser, M., Kombrink, E., Heitz, T., Suh, D.-Y., Legrand, M. and Douglas, C.J. (2010) *LAP6/POLYKETIDE SYNTHASE A* and *LAP5/POLYKETIDE SYNTHASE B* Encode Hydroxyalkyl α-Pyrone Synthases Required

for Pollen Development and Sporopollenin Biosynthesis in *Arabidopsis thaliana*. *The Plant Cell*, **22**, 4045-4066.

Kocyan, A., Zhang, L.-B., Schaefer, H. and Renner, S.S. (2007) A multi-locus chloroplast phylogeny for the Cucurbitaceae and its implications for character evolution and classification. *Molecular Phylogenetics and Evolution*, **44**, 553-577.

Kon, T. and Yoshikawa, N. (2014) Induction and maintenance of DNA methylation in plant promoter sequences by apple latent spherical virus-induced transcriptional gene silencing. *Frontiers in Microbiology*, **5**.

Korpelainen, H. (1998) Labile sex expression in plants. *Biological Reviews*, **73**, 157-180.

Korwar, A.M., Vannuruswamy, G., Jagadeeshaprasad, M.G., Jayaramaiah, R.H., Bhat, S., Regin, B.S., Ramaswamy, S., Giri, A.P., Mohan, V., Balasubramanyam, M. and Kulkarni, M.J. (2015) Development of Diagnostic Fragment Ion Library for Glycated Peptides of Human Serum Albumin: Targeted Quantification in Prediabetic, Diabetic, and Microalbuminuria Plasma by Parallel Reaction Monitoring, SWATH, and MSE. *Molecular & Cellular Proteomics*, **14**, 2150-2159.

Kotake, T., Hojo, S., Yamaguchi, D., Aohara, T., Konishi, T. and Tsumuraya, Y. (2007) Properties and Physiological Functions of UDP-Sugar Pyrophosphorylase in *Arabidopsis*. *Bioscience, Biotechnology, and Biochemistry*, **71**, 761-771.

Kouonon, L.C., Jacquemart, A.-L., Zoro Bi, A.I., Bertin, P., Baudoin, J.-P. and Dje, Y. (2009) Reproductive biology of the andromonoecious *Cucumis melo* subsp. agrestis (Cucurbitaceae). *Annals of Botany*, **104**, 1129-1139.

Krizek, B.A. and Meyerowitz, E.M. (1996) The *Arabidopsis* homeotic genes *APETALA3* and *PISTILLATA* are sufficient to provide the B class organ identity function. *Development*, **122**, 11.

Kumagai, M.H., Donson, J., della-Cioppa, G., Harvey, D., Hanley, K. and Grill, L.K. (1995) Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. *Proceedings of the National Academy of Sciences*, **92**, 1679.

Kumar, L. and Deodikar, G. (1940) Sex Chromosomes of *Coccinia indica* Wight and Arn. *Current Science*, **9**, 128-130.

Kumar, L. and Viseveshwaraiah, S. (1952) Sex mechanism in *Coccinia indica* Wight and Arn. *Nature*, **170**, 330.

Lago, C., Clerici, E., Dreni, L., Horlow, C., Caporali, E., Colombo, L. and Kater, M.M. (2005) The *Arabidopsis* TFIID factor AtTAF6 controls pollen tube growth. *Developmental Biology*, **285**, 91-100.

Laibach, F. and Kribben, F.A. (1949) The influence of growth regulators on the development of male and female blossoms in a monoecious plant (*Cucumis sativus* L.). *Bot. Ges.*, **62**, 53-55.

Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357.

Law, T.F., Lebel-Hardenack, S. and Grant, S.R. (2002) Silver enhances stamen development in female white campion (*Silene latifolia* [Caryophyllaceae]). *American Journal of Botany*, **89**, 1014-1020.

Lebel-Hardenack, S., Hauser, E., Law, T.F., Schmid, J. and Grant, S.R. (2002) Mapping of Sex Determination Loci on the White Campion (*Silene latifolia*) Y Chromosome Using Amplified Fragment Length Polymorphism. *Genetics*, **160**, 717.

Lengerova, M. and Vyskot, B. (2001) Sex chromatin and nucleolar analyses in *Rumex acetosa L. Protoplasma*, **217**, 147-153.

Leroux, C., Bouton, S., Kiefer-Meyer, M.-C., Fabrice, T.N., Mareck, A., Guénin, S., Fournet, F., Ringli, C., Pelloux, J., Driouich, A., Lerouge, P., Lehner, A. and Mollet, J.-C. (2015) PECTIN METHYLESTERASE48 Is Involved in *Arabidopsis* Pollen Grain Germination. *Plant Physiology*, **167**, 367.

Lewis, D. (1942) The Evolution Of Sex In Flowering Plants. *Biological Reviews*, **17**, 46-67.

Li, B. and Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323.

Li, Z., Huang, S., Liu, S., Pan, J., Zhang, Z., Tao, Q., Shi, Q., Jia, Z., Zhang, W., Chen, H., Si, L., Zhu, L. and Cai, R. (2009) Molecular Isolation of the *M* Gene Suggests That a Conserved-Residue Conversion Induces the Formation of Bisexual Flowers in Cucumber Plants. *Genetics*, **182**, 1381.

Liu, Y., Schiff, M., Marathe, R. and Dinesh-Kumar, S.P. (2002) Tobacco *Rar1*, *EDS1* and *NPR1/NIM1* like genes are required for *N*-mediated resistance to tobacco mosaic virus. *The Plant Journal*, **30**, 415-429.

Liu, Z., Moore, P.H., Ma, H., Ackerman, C.M., Ragiba, M., Yu, Q., Pearl, H.M., Kim, M.S., Charlton, J.W., Stiles, J.I., Zee, F.T., Paterson, A.H. and Ming, R. (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. *Nature*, **427**, 348.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. *Methods*, **25**, 402-408.

Lloyd, D.G. (1975) The maintenance of gynodioecy and androdioecy in angiosperms. *Genetica*, **45**, 325-339.

Lloyd, D.G. (1976) The transmission of genes via pollen and ovules in gynodioecious angiosperms. *Theoretical Population Biology*, **9**, 299-316.

Lloyd, D.G. (1979) Evolution towards dioecy in heterostylous populations. *Plant Systematics and Evolution*, **131**, 71-80.

Ma, H., Moore, P.H., Liu, Z., Kim, M.S., Yu, Q., Fitch, M.M.M., Sekioka, T., Paterson, A.H. and Ming, R. (2004) High-Density Linkage Mapping Revealed Suppression of Recombination at the Sex Determination Locus in Papaya. *Genetics*, **166**, 419.

Manzano, S., Megías, Z., Martínez, C., García, A., Aguado, E., Chileh, T., López-Alonso, D., García-Maroto, F., Kejnovský, E., Široký, J., Kubát, Z., Králová, T., Vyskot, B. and Jamilena, M. (2017) Overexpression of a flower-specific aerolysin-like protein from

the dioecious plant *Rumex acetosa* alters flower development and induces male sterility in transgenic tobacco. *The Plant Journal*, **89**, 58-72.

Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., Pitrat, M., Dogimont, C. and Bendahmane, A. (2009) A transposon-induced epigenetic change leads to sex determination in melon. *Nature*, **461**, 1135.

Matsunaga, S., Hizume, M., Kawano, S. and Kuroiwa, T. (1994) Cytological Analyses in *Melandrium album*: Genome Size, Chromosome Size and Fluorescence *in situ* Hybridization. *CYTOLOGIA*, **59**, 135-141.

Matsunaga, S., Kawano, S., Takano, H., Uchida, H., Sakai, A. and Kuroiwa, T. (1996) Isolation and developmental expression of male reproductive organ-specific genes in a dioecious campion, Melandrium album (Silene latifolia). *The Plant Journal*, **10**, 679-689.

McDaniel, B.K. and Binder, B.M. (2012) Ethylene receptor 1 (ETR1) is sufficient and has the predominant role in mediating inhibition of ethylene responses by silver in *Arabidopsis thaliana*. *Journal of Biological Chemistry*, **287**, 26094-26103.

Ming, R., Bendahmane, A. and Renner, S.S. (2011) Sex Chromosomes in Land Plants. *Annual Review of Plant Biology*, **62**, 485-514.

Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J.H., Senin, P., Wang, W., Ly, B.V., Lewis, K.L.T., Salzberg, S.L., Feng, L., Jones, M.R., Skelton, R.L., Murray, J.E., Chen, C., Qian, W., Shen, J., Du, P., Eustice, M., Tong, E., Tang, H., Lyons, E., Paull, R.E., Michael, T.P., Wall, K., Rice, D.W., Albert, H., Wang, M.-L., Zhu, Y.J., Schatz, M., Nagarajan, N., Acob, R.A., Guan, P., Blas, A., Wai, C.M., Ackerman, C.M., Ren, Y., Liu, C., Wang, J., Wang, J., Na, J.-K., Shakirov, E.V., Haas, B., Thimmapuram, J., Nelson, D., Wang, X., Bowers, J.E., Gschwend, A.R., Delcher, A.L., Singh, R., Suzuki, J.Y., Tripathi, S., Neupane, K., Wei, H., Irikura, B., Paidi, M., Jiang, N., Zhang, W., Presting, G., Windsor, A., Navajas-Pérez, R., Torres, M.J., Feltus, F.A., Porter, B., Li, Y., Burroughs, A.M., Luo, M.-C., Liu, L., Christopher, D.A., Mount, S.M., Moore, P.H., Sugimura, T., Jiang, J., Schuler, M.A., Friedman, V., Mitchell-Olds, T., Shippen, D.E., dePamphilis, C.W., Palmer, J.D., Freeling, M., Paterson, A.H., Gonsalves, D., Wang, L. and Alam, M. (2008) The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). *Nature*, **452**, 991.

Mizuno, S., Osakabe, Y., Maruyama, K., Ito, T., Osakabe, K., Sato, T., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2007) Receptor-like protein kinase 2 (RPK 2) is a novel factor controlling anther development in *Arabidopsis thaliana*. *The Plant Journal*, **50**, 751-766.

Mohanty, J.N., Nayak, S., Jha, S. and Joshi, R.K. (2017) Transcriptome profiling of the floral buds and discovery of genes related to sex-differentiation in the dioecious cucurbit *Coccinia grandis* (L.) Voigt. *Gene*, **626**, 395-406.

Moore, R.C., Kozyreva, O., Lebel-Hardenack, S., Siroky, J., Hobza, R., Vyskot, B. and Grant, S.R. (2003) Genetic and Functional Analysis of *DD44*, a Sex-Linked Gene From the Dioecious Plant *Silene latifolia*, Provides Clues to Early Events in Sex Chromosome Evolution. *Genetics*, **163**, 321.

Morant, M., Jørgensen, K., Schaller, H., Pinot, F., Møller, B.L., Werck-Reichhart, D. and Bak, S. (2007) CYP703 Is an Ancient Cytochrome P450 in Land Plants Catalyzing in-Chain Hydroxylation of Lauric Acid to Provide Building Blocks for Sporopollenin Synthesis in Pollen. *The Plant Cell*, **19**, 1473.

Muyle, A., Käfer, J., Zemp, N., Mousset, S., Picard, F. and Marais, G.A.B. (2016) SEX-DETector: A Probabilistic Approach to Study Sex Chromosomes in Non-Model Organisms. *Genome Biology and Evolution*, **8**, 2530-2543.

Muyle, A., Zemp, N., Deschamps, C., Mousset, S., Widmer, A. and Marais, G.A. (2012) Rapid *de novo* evolution of X chromosome dosage compensation in *Silene latifolia*, a plant with young sex chromosomes. *PLoS biology*, **10**, e1001308.

Nasrallah, J.B. and Nasrallah, M.E. (1993) Pollen[mdash]Stigma Signaling in the Sporophytic Self-Incompatibility Response. *The Plant Cell*, **5**, 1325.

Navajas-Pérez, R., de la Herrán, R., López González, G., Jamilena, M., Lozano, R., Ruiz Rejón, C., Ruiz Rejón, M. and Garrido-Ramos, M.A. (2005) The Evolution of Reproductive Systems and Sex-Determining Mechanisms Within Rumex (Polygonaceae) Inferred from Nuclear and Chloroplastidial Sequence Data. *Molecular Biology and Evolution*, **22**, 1929-1939.

Newbigin, E., Anderson, M.A. and Clarke, A.E. (1993) Gametophytic Self-Incompatibility Systems. *The Plant Cell*, **5**, 1315.

O'Malley, R.C., Rodriguez, F.I., Esch, J.J., Binder, B.M., O'Donnell, P., Klee, H.J. and Bleecker, A.B. (2005) Ethylene-binding activity, gene expression levels, and receptor system output for ethylene receptor family members from *Arabidopsis* and tomato. *The Plant Journal*, **41**, 651-659.

Owens, K.W., Peterson, C.E. and Tolla, G.E. (1980) Induction of perfect flowers on gynoecious muskmelon by silver nitrate and aminoethoxyvinylglycine. *Report, Cucurbit Genetics Cooperative*, 35-36.

Pagnussat, G.C., Yu, H.-J., Ngo, Q.A., Rajani, S., Mayalagu, S., Johnson, C.S., Capron, A., Xie, L.-F., Ye, D. and Sundaresan, V. (2005) Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. *Development*, **132**, 603.

Painter, T.S. (1921) THE Y-CHROMOSOME IN MAMMALS. *Science*, **53**, 503.

Park, J.-H., Ishikawa, Y., Yoshida, R., Kanno, A. and Kameya, T. (2003) Expression of AODEF, a B-functional MADS-box gene, in stamens and inner tepals of the dioecious species Asparagus officinalis L. *Plant Molecular Biology*, **51**, 867-875.

Park, J.-I., Ishimizu, T., Suwabe, K., Sudo, K., Masuko, H., Hakozaki, H., Nou, I.-S., Suzuki, G. and Watanabe, M. (2010) UDP-Glucose Pyrophosphorylase is Rate Limiting in Vegetative and Reproductive Phases in *Arabidopsis thaliana*. *Plant and Cell Physiology*, **51**, 981-996.

Parker, J.S. and Clark, M.S. (1991) Dosage sex-chromosome systems in plants. *Plant Science*, **80**, 79-92.

Peele, C., Jordan, C.V., Muangsan, N., Turnage, M., Egelkrout, E., Eagle, P., Hanley-Bowdoin, L. and Robertson, D. (2001) Silencing of a meristematic gene using geminivirus-derived vectors. *The Plant Journal*, **27**, 357-366.

Peng, J., Ilarslan, H., Wurtele, E.S. and Bassham, D.C. (2011) *AtRabD2b* and *AtRabD2c* have overlapping functions in pollen development and pollen tube growth. *BMC Plant Biology*, **11**, 25.

Pfent, C., Pobursky, K.J., Sather, D.N. and Golenberg, E.M. (2005) Characterization of *SpAPETALA3* and *SpPISTILLATA*, B class floral identity genes in *Spinacia oleracea*, and their relationship to sexual dimorphism. *Development Genes and Evolution*, **215**, 132-142.

Pflieger, S., Richard, M.M., Blanchet, S., Meziadi, C. and Geffroy, V. (2013) VIGS technology: an attractive tool for functional genomics studies in legumes. *Functional plant biology*, **40**, 1234-1248.

Qin, Y., Leydon, A.R., Manziello, A., Pandey, R., Mount, D., Denic, S., Vasic, B., Johnson, M.A. and Palanivelu, R. (2009) Penetration of the Stigma and Style Elicits a Novel Transcriptome in Pollen Tubes, Pointing to Genes Critical for Growth in a Pistil. *PLOS Genetics*, **5**, e1000621.

Raju, S., Chithakari, R. and Mustafa, M. (2014) In vitro plantlet regeneration from tendril explants of spine gourd (*Momordica dioica* Roxb. ex. wild). *Indian Journal Of Applied Research*, **4**, 50-52.

Ratcliff, F., Harrison, B.D. and Baulcombe, D.C. (1997) A Similarity Between Viral Defense and Gene Silencing in Plants. *Science*, **276**, 1558.

Ratcliff, F., Martin-Hernandez, A.M. and Baulcombe, D.C. (2001) Technical Advance: Tobacco rattle virus as a vector for analysis of gene function by silencing. *The Plant Journal*, **25**, 237-245.

Rejón, C.R., Jamilena, M., Garrido Ramos, M., Parker, J.S. and Ruiz Rejón, M. (1994) Cytogenetic and molecular analysis of the multiple sex chromosome system of Rumex acetosa. *Heredity*, **72**, 209.

Renner, S.S. (2014) The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, **101**, 1588-1596.

Renner, S.S. and Ricklefs, R.E. (1995) Dioecy and its correlates in the flowering plants. *American Journal of Botany*, **82**, 596-606.

Rhee, S.Y., Osborne, E., Poindexter, P.D. and Somerville, C.R. (2003) Microspore Separation in the *quartet 3* Mutants of *Arabidopsis* Is Impaired by a Defect in a Developmentally Regulated Polygalacturonase Required for Pollen Mother Cell Wall Degradation. *Plant Physiology*, **133**, 1170-1180.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139-140.

Ross, M.D. (1978) The Evolution Of Gynodioecy And Subdioecy. *Evolution*, **32**, 174-188.

Ruiz, M.T., Voinnet, O. and Baulcombe, D.C. (1998) Initiation and Maintenance of Virus-Induced Gene Silencing. *The Plant Cell*, **10**, 937.

Saito, S., Fujii, N., Miyazawa, Y., Yamasaki, S., Matsuura, S., Mizusawa, H., Fujita, Y. and Takahashi, H. (2007) Correlation between development of female flower buds and expression of the *CS-ACS2* gene in cucumber plants. *Journal of Experimental Botany*, **58**, 2897-2907.

Sarath, G. and Mohan Ram, H.Y. (1979) Comparative effect of silver ion and gibberellic acid on the induction of male flowers on female *Cannabis* plants. *Experientia*, **35**, 333-334.

Sasaki, S., Yamagishi, N. and Yoshikawa, N. (2011) Efficient virus-induced gene silencing in apple, pear and Japanese pear using Apple latent spherical virus vectors. *Plant Methods*, **7**, 15.

Sather, D.N., Jovanovic, M. and Golenberg, E.M. (2010) Functional analysis of B and C class floral organ genes in spinach demonstrates their role in sexual dimorphism. *BMC Plant Biology*, **10**, 46.

Schilmiller, A.L., Stout, J., Weng, J.-K., Humphreys, J., Ruegger, M.O. and Chapple, C. (2009) Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in *Arabidopsis*. *The Plant Journal*, **60**, 771-782.

Scutt, C.P., Jenkins, T., Furuya, M. and Gilmartin, P.M. (2002) Male Specific Genes from Dioecious White Campion Identified by Fluorescent Differential Display. *Plant and Cell Physiology*, **43**, 563-572.

Senthil-Kumar, M. and Mysore, K.S. (2011) New dimensions for VIGS in plant functional genomics. *Trends in Plant Science*, **16**, 656-665.

Senthil-Kumar, M. and Mysore, K.S. (2014) Tobacco rattle virus–based virus-induced gene silencing in *Nicotiana benthamiana*. *Nature Protocols*, **9**, 1549.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, **13**, 2498-2504.

Shibata, F., Hizume, M. and Kuroki, Y. (1999) Chromosome painting of Y chromosomes and isolation of a Y chromosome-specific repetitive sequence in the dioecious plant Rumex acetosa. *Chromosoma*, **108**, 266-270.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V. and Zdobnov, E.M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, **31**, 3210-3212.

Smalle, J., Kurepa, J., Yang, P., Emborg, T.J., Babiychuk, E., Kushnir, S. and Vierstra, R.D. (2003) The Pleiotropic Role of the 26S Proteasome Subunit RPN10 in *Arabidopsis* Growth and Development Supports a Substrate-Specific Function in Abscisic Acid Signaling. *The Plant Cell*, **15**, 965-980.

Smith-Unna, R., Boursnell, C., Patro, R., Hibberd, J.M. and Kelly, S. (2016) TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome Research*, **26**, 1134-1144.

Sorensen, A.M., Kröber, S., Unte, U.S., Huijser, P., Dekker, K. and Saedler, H. (2003) The Arabidopsis *ABORTED MICROSPORES* (*AMS*) gene encodes a MYC class transcription factor. *The Plant Journal*, **33**, 413-423.

Sousa, A., Fuchs, J. and Renner, S.S. (2013) Molecular Cytogenetics (FISH, GISH) of *Coccinia grandis*: A ca. 3 myr-Old Species of Cucurbitaceae with the Largest Y/Autosome Divergence in Flowering Plants. *Cytogenetic and Genome Research*, **139**, 107-118.

Spitzer, C., Reyes, F.C., Buono, R., Sliwinski, M.K., Haas, T.J. and Otegui, M.S. (2009) The ESCRT-Related CHMP1A and B Proteins Mediate Multivesicular Body Sorting of Auxin Carriers in *Arabidopsis* and Are Required for Plant Development. *The Plant Cell*, **21**, 749.

Stankovic, L., Stevanovic, D. and Zdravkovic, M. (2005) Some effects of ethrel on plants from Cucurbitaceae [family]. *Arhiv za poljoprivredne nauke*.

Storey, W.B. (1953) Genetics of the papaya. *Journal of Heredity*, **44**, 70-78.

Switzenberg, J.A., Little, H.A., Hammar, S.A. and Grumet, R. (2014) Floral primordia-targeted ACS (1-aminocyclopropane-1-carboxylate synthase) expression in transgenic *Cucumis melo* implicates fine tuning of ethylene production mediating unisexual flower development. *Planta*, **240**, 797-808.

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. and von Mering, C. (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Research*, **45**, D362-D368.

Tao, Q., Niu, H., Wang, Z., Zhang, W., Wang, H., Wang, S., Zhang, X. and Li, Z. (2018) Ethylene responsive factor ERF110 mediates ethylene-regulated transcription of a sex determination-related orthologous gene in two *Cucumis* species. *Journal of Experimental Botany*, **69**, 2953-2965.

Tao, X. and Zhou, X. (2008) Pathogenicity of a naturally occurring recombinant DNA satellite associated with tomato yellow leaf curl China virus. *Journal of General Virology*, **89**, 306-311.

Thiruvengadam, M. and Chung, I.-M. (2011) Establishment of an efficient *Agrobacterium tumefaciens*-mediated leaf disc transformation of spine gourd (*Momordica dioica* Roxb. ex Willd). *African Journal of Biotechnology*, **10**, 19337-19345.

Tian, G.-W., Chen, M.-H., Zaltsman, A. and Citovsky, V. (2006) Pollen-specific pectin methylesterase involved in pollen tube growth. *Developmental Biology*, **294**, 83-91.

Toujani, W., Muñoz-Bertomeu, J., Flores-Tornero, M., Rosa-Téllez, S., Anoman, A.D., Alseekh, S., Fernie, A.R. and Ros, R. (2013) Functional Characterization of the Plastidial 3-Phosphoglycerate Dehydrogenase Family in *Arabidopsis*. *Plant Physiology*, **163**, 1164-1178.

Tsugama, D., Matsuyama, K., Ide, M., Hayashi, M., Fujino, K. and Masuda, K. (2017) A putative MYB35 ortholog is a candidate for the sex-determining genes in *Asparagus officinalis*. *Scientific Reports*, **7**, 41497.

Uchida, W., Matsunaga, S., Sugiyama, R., Kazama, Y. and Kawano, S. (2003) Morphological development of anthers induced by the dimorphic smut fungus *Microbotryum violaceum* in female flowers of the dioecious plant *Silene latifolia*. *Planta*, **218**, 240-248.

Urasaki, N., Tarora, K., Shudo, A., Ueno, H., Tamaki, M., Miyagi, N., Adaniya, S. and Matsumura, H. (2012) Digital Transcriptome Analysis of Putative Sex-Determination Genes in Papaya (*Carica papaya*). *PLOS ONE*, **7**, e40904.

Van Damme, D., Coutuer, S., De Rycke, R., Bouget, F.-Y., Inzé, D. and Geelen, D. (2006) Somatic Cytokinesis and Pollen Maturation in *Arabidopsis* Depend on TPLATE, Which Has Domains Similar to Coat Proteins. *The Plant Cell*, **18**, 3502-3518.

Varkonyi-Gasic, E., Wu, R., Wood, M., Walton, E.F. and Hellens, R.P. (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods*, **3**, 12.

Volz, S.M. and Renner, S.S. (2008) Hybridization, polyploidy, and evolutionary transitions between monoecy and dioecy in Bryonia (Cucurbitaceae). *American Journal of Botany*, **95**, 1297-1306.

Vyskot, B. and Hobza, R. (2004) Gender in plants: sex chromosomes are emerging from the fog. *Trends in Genetics*, **20**, 432-438.

Wang, L., Wang, W., Wang, Y.-Q., Liu, Y.-Y., Wang, J.-X., Zhang, X.-Q., Ye, D. and Chen, L.-Q. (2013) *Arabidopsis* Galacturonosyltransferase (GAUT) 13 and GAUT14 Have Redundant Functions in Pollen Tube Growth. *Molecular Plant*, **6**, 1131-1148.

Waterhouse, P.M. and Fusaro, A.F. (2006) Viruses Face a Double Defense by Plant Small RNAs. *Science*, **313**, 54.

Westergaard, M. (1940) Studies on cytology and sex determination in polyploid forms of Melandrium album. *Danks. Botan. Arkiv*, **10**, 1-31.

Westergaard, M. (1946) Aberrant Y Chromosomes And Sex Expression In *Melandrium Album*. *Hereditas*, **32**, 419-443.

Westergaard, M. (1958) The Mechanism of Sex Determination in Dioecious Flowering Plants. In *Advances in Genetics* (Demerec, M. ed: Academic Press, pp. 217-281.

Wittwer, S.H. and Bukovac, M.J. (1958) The Effects of Gibberellin on Economic Crops. *Economic Botany*, **12**, 213-255.

WU, C., JIA, L. and GOGGIN, F. (2011) The reliability of virus-induced gene silencing experiments using tobacco rattle virus in tomato is influenced by the size of the vector control. *Molecular Plant Pathology*, **12**, 299-305.

Wu, T., Qin, Z., Zhou, X., Feng, Z. and Du, Y. (2010) Transcriptome profile analysis of floral sex determination in cucumber. *Journal of Plant Physiology*, **167**, 905-913.

Wuest, S.E., O'Maoileidigh, D.S., Rae, L., Kwasniewska, K., Raganelli, A., Hanczaryk, K., Lohan, A.J., Loftus, B., Graciet, E. and Wellmer, F. (2012) Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proceedings of the National Academy of Sciences*, **109**, 13452.

Xie, B., Wang, X. and Hong, Z. (2010) Precocious pollen germination in *Arabidopsis* plants with altered callose deposition during microsporogenesis. *Planta*, **231**, 809-823.

Xiong, S.X., Lu, J.Y., Lou, Y., Teng, X.D., Gu, J.N., Zhang, C., Shi, Q.S., Yang, Z.N. and Zhu, J. (2016) The transcription factors MS188 and AMS form a complex to activate the expression of *CYP703A2* for sporopollenin biosynthesis in *Arabidopsis thaliana*. *The Plant Journal*, **88**, 936-946.

Xu, J., Yang, C., Yuan, Z., Zhang, D., Gondwe, M.Y., Ding, Z., Liang, W., Zhang, D. and Wilson, Z.A. (2010) The *ABORTED MICROSPORES* Regulatory Network Is Required for Postmeiotic Male Reproductive Development in *Arabidopsis thaliana*. *The Plant Cell*, **22**, 91-107.

Yampolsky, C. (1922) Distribution of sex forms in the phanerogamic flora. *Bibliotheca Genetica*, **3**, 1-62.

Yang, C., Vizcay-Barrena, G., Conner, K. and Wilson, Z.A. (2007) MALE STERILITY1 Is Required for Tapetal Development and Pollen Wall Biosynthesis. *The Plant Cell*, **19**, 3530.

Yang, K., Wang, H., Xue, S., Qu, X., Zou, J. and Le, J. (2014) Requirement for A-type cyclin-dependent kinase and cyclins for the terminal division in the stomatal lineage of *Arabidopsis*. *Journal of Experimental Botany*, **65**, 2449-2461.

Yang, L., Gong, F., Xiong, E. and Wang, W. (2015) Proteomics: a promising tool for research on sex-related differences in dioecious plants. *Frontiers in Plant Science*, **6**.

Yang, X., Makaroff, C.A. and Ma, H. (2003) The *Arabidopsis MALE MEIOCYTE DEATH1* Gene Encodes a PHD-Finger Protein That Is Required for Male Meiosis. *The Plant Cell*, **15**, 1281.

Yang, Y. and Karlson, D.T. (2011) Overexpression of *AtCSP4* affects late stages of embryo development in *Arabidopsis*. *Journal of Experimental Botany*, **62**, 2079-2091.

Yin, T., DiFazio, S.P., Gunter, L.E., Zhang, X., Sewell, M.M., Woolbright, S.A., Allan, G.J., Kelleher, C.T., Douglas, C.J., Wang, M. and Tuskan, G.A. (2008) Genome structure and emerging evidence of an incipient sex chromosome in *Populus*. *Genome Research*, **18**, 422-430.

Yin, T. and Quinn, J.A. (1995) Tests of a mechanistic model of one hormone regulating both sexes in *Cucumis sativus* (Cucurbitaceae). *American Journal of Botany*, 1537-1546.

Yu, Q., Navajas-Pérez, R., Tong, E., Robertson, J., Moore, P.H., Paterson, A.H. and Ming, R. (2008) Recent Origin of Dioecious and Gynodioecious Y Chromosomes in Papaya. *Tropical Plant Biology*, **1**, 49-57.

Zhan, H., Zhong, Y., Yang, Z. and Xia, H. (2015) Enzyme activities of *Arabidopsis* inositol polyphosphate kinases AtIPK2α and AtIPK2β are involved in pollen development, pollen tube guidance and embryogenesis. *The Plant Journal*, **82**, 758-771.

Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., Shi, Z., Chambers, M.C., Zimmerman, L.J., Shaddox, K.F., Kim, S., Davies, S.R., Wang, S., Wang, P., Kinsinger, C.R., Rivers, R.C., Rodriguez, H., Townsend, R.R., Ellis, M.J.C., Carr, S.A., Tabb, D.L.,

Coffey, R.J., Slebos, R.J.C., Liebler, D.C. and the, N.C. (2014a) Proteogenomic characterization of human colon and rectal cancer. *Nature*, **513**, 382.

Zhang, C., Bradshaw, J.D., Whitham, S.A. and Hill, J.H. (2010) The Development of an Efficient Multipurpose Bean Pod Mottle Virus Viral Vector Set for Foreign Gene Expression and RNA Silencing. *Plant Physiology*, **153**, 52.

Zhang, D., Liu, D., Lv, X., Wang, Y., Xun, Z., Liu, Z., Li, F. and Lu, H. (2014b) The Cysteine Protease CEP1, a Key Executor Involved in Tapetal Programmed Cell Death, Regulates Pollen Development in *Arabidopsis*. *The Plant Cell*, **26**, 2939-2961.

Zhang, J., Boualem, A., Bendahmane, A. and Ming, R. (2014c) Genomics of sex determination. *Current Opinion in Plant Biology*, **18**, 110-116.

Zhao, D.-Z., Wang, G.-F., Speal, B. and Ma, H. (2002) The *EXCESS MICROSPOROCYTES1* gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the *Arabidopsis* anther. *Genes & Development*, **16**, 2021-2031.

Zhao, L.-N., Shen, L.-K., Zhang, W.-Z., Zhang, W., Wang, Y. and Wu, W.-H. (2013) Ca^{2+}-Dependent Protein Kinase11 and 24 Modulate the Activity of the Inward Rectifying K^+ Channels in *Arabidopsis* Pollen Tubes. *The Plant Cell*, **25**, 649.

Zhu, E., You, C., Wang, S., Cui, J., Niu, B., Wang, Y., Qi, J., Ma, H. and Chang, F. (2015) The DYT1-interacting proteins bHLH010, bHLH089 and bHLH091 are redundantly required for Arabidopsis anther development and transcriptome. *The Plant Journal*, **83**, 976-990.

Zhu, J., Wu, X., Yuan, S., Qian, D., Nan, Q., An, L. and Xiang, Y. (2014) Annexin5 Plays a Vital Role in *Arabidopsis* Pollen Development via Ca2+-Dependent Membrane Trafficking. *PLOS ONE*, **9**, e102407.

Zluvova, J., Nicolas, M., Berger, A., Negrutiu, I. and Monéger, F. (2006) Premature arrest of the male flower meristem precedes sexual dimorphism in the dioecious plant *Silene latifolia*. *Proceedings of the National Academy of Sciences*, **103**, 18854.

AUTHOR'S PUBLICATIONS

CrossMark

De novo transcriptome assembly from flower buds of dioecious, gynomonoecious and chemically masculinized female *Coccinia grandis* reveals genes associated with sex expression and modification

Ravi Suresh Devani[1], Sangram Sinha[2], Jayeeta Banerjee[1], Rabindra Kumar Sinha[2], Abdelhafid Bendahmane[3] and Anjan Kumar Banerjee[1*] iD

Abstract

Background: *Coccinia grandis* (ivy gourd), is a dioecious member of Cucurbitaceae having heteromorphic sex chromosomes. Chromosome constitution of male and female plants of *C. grandis* is 22A + XY and 22A + XX respectively. Earlier we showed that a unique gynomonoecious form of *C. grandis* (22A + XX) also exists in nature bearing morphologically hermaphrodite flowers (GyM-H). Additionally, application of silver nitrate (AgNO$_3$) on female plants induces stamen development leading to the formation of morphologically hermaphrodite flowers (Ag-H) despite the absence of Y-chromosome. Due to the unavailability of genome sequence and the slow pace at which sex-linked genes are identified, sex expression and modification in *C. grandis* are not well understood.

Results: We have carried out a comprehensive RNA-Seq study from early-staged male, female, GyM-H, and Ag-H as well as middle-staged male and GyM-H flower buds. A de novo transcriptome was assembled using Trinity and annotated by BLAST2GO and Trinotate pipelines. The assembled transcriptome consisted of 467,233 'Trinity Transcripts' clustering into 378,860 'Trinity Genes'. Female_Early_vs_Male_Early, Ag_Early_vs_Female_Early, and GyM-H_Middle_vs_Male_Middle comparisons exhibited 35,694, 3574, and 14,954 differentially expressed transcripts respectively. Further, qRT-PCR analysis of selected candidate genes validated digital gene expression profiling results. Interestingly, ethylene response-related genes were found to be upregulated in female buds compared to male buds. Also, we observed that AgNO$_3$ treatment suppressed ethylene responses in Ag-H flowers by downregulation of ethylene-responsive transcription factors leading to stamen development. Further, GO terms related to stamen development were enriched in early-staged male, GyM-H, and Ag-H buds compared to female buds supporting the fact that stamen growth gets arrested in female flowers.

Conclusions: Suppression of ethylene responses in both male and Ag-H compared to female buds suggests a probable role of ethylene in stamen suppression similar to monoecious cucurbits such as melon and cucumber. Also, pollen fertility associated GO terms were depleted in middle-staged GyM-H buds compared to male buds indicating the necessity of Y-chromosome for pollen fertility. Overall, this study would enable identification of new sex-biased genes for further investigation of stamen arrest, pollen fertility, and AgNO$_3$-mediated sex modification.

Keywords: *Coccinia grandis*, Dioecious, Gynomonoecious, De novo transcriptome, Stamen arrest, Silver nitrate, Ethylene, Pollen fertility

---

* Correspondence: akb@iiserpune.ac.in
[1]Biology Division, Indian Institute of Science Education and Research (IISER), Pune, Pune, Maharashtra, India
Full list of author information is available at the end of the article

Devani et al. BMC Plant Biology (2017) 17:241

Page 2 of 15

Background

Monoecy, dioecy, and hermaphroditism are the three major sexual forms observed among the flowering plants. Ninety (90%) of angiosperms are found to be hermaphrodite (both male and female organs are in the same flower), while 5% plant species exhibit monoecy (male and female flowers are on the same plant) and remaining 5% show dioecy (male and female flowers are in separate plant) [1]. Dioecism provides a unique opportunity to study the genetic basis of sex determination. *Silene latifolia* (Caryophyllaceae), *Rumex acetosa* (Polygonaceae), *Carica papaya* (Caricaceae), *Spinacia oleracea* (Chenopodiaceae) and *Populus* (Salicaceae), have been well characterized to understand the mechanism of sex determination [2–5]. However, the molecular mechanism and the genes that govern sex determination are not well understood.

Coccinia grandis (L.) Voigt, a dioecious member of Cucurbitaceae family having an inferior ovary has received comparatively less attention. Members of Cucurbitaceae family exhibit variety of sexual forms [6]. Apart from its rich medicinal value, *C. grandis*, commonly known as ivy gourd, is also used as a vegetable. *Coccinia grandis* bears male and female unisexual flowers on separate plants. Similar to *Silene latifolia* (Caryophyllaceae), the sex in *Coccinia grandis* is determined by the presence of Y-chromosome [7–9]. The chromosome constitution of male and female plants is 22A + XY and 22A + XX respectively, where Y-chromosome is larger than the X-chromosome [10–12]. The male flower consists of three convoluted (bithecous) stamens [13, 14] and lacks female reproductive organs; however, the female flower consists of three rudimentary stamens surrounding the three fused carpels with an inferior ovary [14]. There are two ways by which unisexual flower development can be achieved. One of the ways is when both male and female sex organ primordia are initiated at early stages of flower development, but at later stages the opposite sex organs are aborted as in *Silene latifolia* [15]. Another way is that organ primordia of the opposite sex organs do not develop at all as shown in *Thalictrum dioicum* [16]. Also, there are flowers, wherein the inappropriate sex organs are retained in rudimentary form (instead of getting aborted) as in *Rumex* and *C. grandis* [2, 14]. Additionally, *Coccinia grandis* shows sex modification upon application of AgNO$_3$ leading to the development of stamens in female flower (such flower will be referred to as Ag-H) as described in our previous report [14]. Ag$^+$ has been long known to be an inhibitor of ethylene response [17]. It has been suggested that the binding of Ag$^+$ to the ethylene receptor inhibits the conformational change, which maintains the receptor in the active conformation [18]. Application of silver compounds such as silver nitrate (AgNO$_3$) or silver thiosulphate (Ag$_2$S$_2$O$_3$)

masculinizes monoecious plants such as *Cucumis sativus* as well as female plants of dioecious species such as *Silene latifolia* and *Cannabis sativa* [19–21]. However, the mechanism of action by which Ag$^+$ induces stamen development is not known till date [21].

Despite the interesting discovery of sex chromosomes in dioecious plants more than 50 years ago, the mechanism of sex determination remains poorly understood [22, 23]. This is primarily because of the slow pace at which sex-linked genes were identified from dioecious species (one to two genes/year) [24]. However, the improvement in NGS technology has already started changing the situation by accelerating the rate of sex-linked gene identification. The NGS-based approach has a big advantage that it does not require prior knowledge of the gene sequences to be investigated. Recently, an NGS-based RNA-Seq approach was applied to *Silene latifolia*, which was the first report that demonstrated the phenomenon of dosage compensation in plants [24]. A comparative transcriptomics approach was applied to papaya, a trioecious species, to identify the candidate genes for sex determination. This study led to the identification of 312 unique tags that were specifically mapped to the primitive sex chromosome (X or Y^h) sequences in papaya [5]. A genome-wide transcriptional profiling of apical tissue of a gynoecious mutant (Csg-G) and the monoecious wild-type (Csg-M) of cucumber was also performed to isolate genes involved in sex determination. This study revealed that genes involved in plant hormone signaling pathways, such as *ACS*, *Asr1*, *CsIAA2*, *CS-AUX1*, and *TLP*, and their crosstalk might play a critical role in the sex determination. Authors have also predicted the regulation of some transcription factors, including *EREBP*-9, in sex determination [25]. In another study, transcriptome sequencing was carried out from cucumber flower buds of two near-isogenic lines, WI1983G, a gynoecious plant which bears only pistillate flowers and WI1983H, a hermaphroditic plant which bears only bisexual flowers [26]. This study identified differentially expressed genes as well as putative SSR and SNP markers between flowers of two different sexes. T Akagi, IM Henry, R Tao and L Comai [27] sequenced genomic DNA, mRNA as well as small RNA from flower buds of persimmon and identified a Y-chromosome–encoded small RNA, *OGI*, that targets a homeodomain transcription factor *MeGI* regulating pollen fertility in a dosage-dependent manner. A recent de novo transcriptomics study in garden *Asparagus* identified 570 differentially expressed genes, where genes involved in pollen microspore and tapetum development were shown to be specifically expressed in males and supermales in contrast to females [28].

In order to identify sex-biased genes in *C. grandis* and to elucidate the mechanism of AgNO$_3$ mediated sex modification, a comprehensive RNA-Seq study from early-staged male (M), female (F), GyM-H and Ag-H as

Devani *et al. BMC Plant Biology* (2017) 17:241

Page 3 of 15

well as middle-staged male and GyM-H flower buds was carried out. De novo transcriptome was assembled to identify *C. grandis* homologs of various flower development genes. Digital expression profiling was undertaken to identify sex-biased genes which might play a pivotal role in the arrest of stamen development in female flowers, genes that promote anther development in female flowers upon AgNO₃ treatment and genes controlling pollen fertility in male flowers.

Methods

Flower bud collection and RNA isolation

Clones of wild-type male, female and gynomonoecious (GyM) forms of *C. grandis* were grown in the experimental plot at IISER Pune, India. Gynomonoecious (GyM) plant bears pistillate flowers (GyM-F) as well as morphologically hermaphrodite flowers (GyM-H) (Herbarium Voucher: Tripura University Campus, Karmakar, 433). Foliar spray of 35 mM AgNO₃ solution to the basal leaves on some of the female plants led to the development of morphologically hermaphrodite flowers (Ag-H) as per our earlier observation [14]. Ag-H flower buds were morphologically similar to GyM-H flowers. Flower buds from male (M), female (F), GyM-H and Ag-H were harvested separately in liquid nitrogen and categorized into early and middle stages based on our previous study [14] (Fig. 1). In early-staged male flower buds, only stamens are present with no sign of carpel initials. Whereas, early-staged female flowers (stages 3–4, Fig. 1) have both carpel and stamen primordia. Stamen growth in female flowers gets arrested around stages 4–5. In the hermaphrodite flowers of gynomonoecious plant, however, both stamens and carpels develop simultaneously during early as well as

Fig. 1 Different stages of *C. grandis* flower buds selected for RNA-Seq analysis. (**a**) Male, (**b**) Female flower buds. GyM-H and Ag-H buds sized similar to female buds were chosen for RNA seq study as described in our previous work (Ghadge et al., 2014). Scale bar = 1 cm

middle stages of development. Our selection of early-staged flower buds was carried out such that the event of stamen inhibition in female flowers can be analysed. Whereas, middle-staged flower buds were chosen such that meiosis-stage and pollen maturation event can be investigated. Total RNA was isolated by TRIzol reagent (Invitrogen) following the manufacturer's instructions. RNA quality was assessed using an Agilent Bioanalyzer RNA nanochip, and RNA samples with RIN > 8.0 were used for library preparation.

RNA-Seq library preparation and sequencing

Library preparation was performed at Genotypic Technology's Genomics facility, Bangalore using Illumina TruSeq RNA Sample Preparation Kit according to the manufacturer's specifications. RNA sequencing libraries were prepared in duplicate for early-staged male (M), female (F), GyM-H and Ag-H flower buds, as well as middle-staged male (M) and GyM-H flower buds. The quality of all the twelve libraries and insert size distribution was assessed using an Agilent High Sensitivity Bioanalyzer Chip. Libraries showed a peak in the range of 250–1000 bp. The effective sequencing insert size was 130–880 bp, and the inserts were flanked by adaptors whose combined size was 120 bp. Libraries were quantified using Qubit and sequenced on Illumina NextSeq 500 platform, producing 2 X 150-nucleotide paired-end reads. RNA-Seq data generated in this study has been deposited in the NCBI SRA study SRP111347.

Pre-processing of Illumina reads and de novo transcriptome assembly

Raw RNA-Seq reads were processed using Trimmomatic v0.33 for trimming adapters as well as low-quality bases from ends of the reads [29]. Poor quality reads with average Phred quality score < 20 and reads with length < 36 were also filtered out. The resulting set of good quality reads were then assembled with Trinity v2.1.1 software using default parameters [30, 31].

The quality of the resulting assembly was assessed by various methods. First of all, RNA-Seq read representation of the assembly was checked using bowtie2 [32]. Ex90N50 transcript contig length (the contig N50 value based on the set of transcripts representing 90% of the expression data) was computed using *contig_ExN50_statistic.pl* script bundled with Trinity. Then the representation of full-length reconstructed protein-coding genes was studied. The assembled transcripts were compared with Swiss-Prot using BLAST and the hits were analyzed using a perl script *blast_outfmt6_group_segments.tophit_coverage.pl*, provided with the Trinity package. BUSCO (Benchmarking Universal Single-Copy Orthologs) was used to explore completeness of the transcriptome according to conserved ortholog content [33]. Finally, TransRate was used to
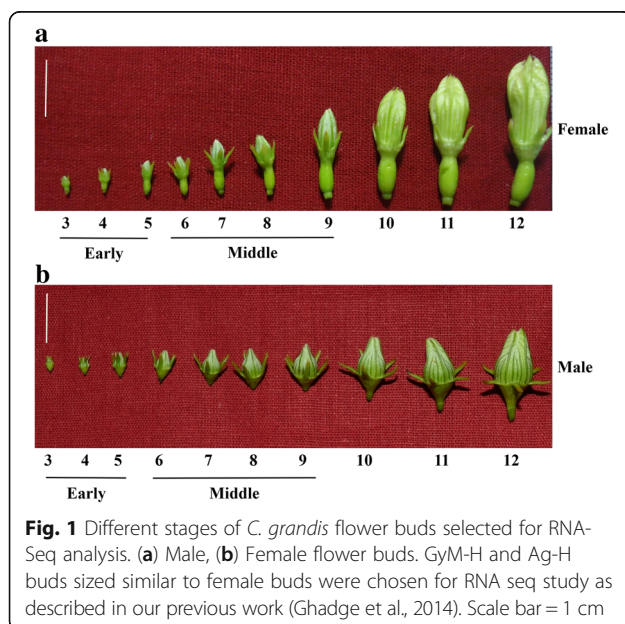
compare the assembly to the publicly available *Cucumis sativus* protein-coding primary transcript sequences [34].

Annotation of the de novo-assembled transcripts

The de novo-assembled transcripts were compared with the viridiplantae sequences from nr and Swiss-Prot database using BLASTX with an e-value threshold of 1e-3 [35]. BLAST output generated from this comparison was loaded into BLAST2GO for mapping GO terms to the transcripts and annotation [36]. Enzyme codes and KEGG pathway mapping were also carried out. ANNEX (Annotation Expander) was used to enhance the annotations. Finally, GO-Slim mapping was applied to get a broad overview of the ontology content.

In addition to BLAST2GO facilitated annotation, Trinotate pipeline was used to carry out comprehensive functional annotation of the transcripts leveraging various annotation databases (eggNOG/GO/KEGG databases) [31]. Trinotate pipeline also included identification of open reading frames, homology search against Swiss-Prot and TrEMBL. Protein domain identification was carried out using HMMER/PFAM. Protein signal peptide and transmembrane domains were predicted by signalP and tmHMM respectively.

Transcript quantification and differential expression analysis

align_and_estimate_abundance.pl script from Trinity package was applied to align cleaned reads from each library to the de novo transcriptome using bowtie and to estimate the transcript abundance using RSEM [37]. *abundance_estimates_to_matrix.pl* script was used to construct a matrix of counts and a matrix of normalized expression values. *PtR* script was used to generate correlation matrix and Principal Component Analysis (PCA) plot for comparing replicates across all the samples. Differential expression analysis was carried out with two biological replicates from the count matrix using *run_DE_analysis.pl* with edgeR as the method of choice [38]. *analyze_diff_expr.pl* script was used to examine GO enrichment and to extract all transcripts that had *p*-values at most 1e-3 and were at least 2^2 fold differentially expressed. The DE features were partitioned into clusters with similar expression patterns by *define_clusters_by_cutting_tree.pl* script with Ptree method.

Validation of differentially expressed genes by qRT-PCR

For expression analysis, qRT-PCR was carried out using aliquots of the same RNA samples that were used for RNA sequencing. Two micrograms (2 μg) of total RNA was used for complementary DNA (cDNA) synthesis by SuperScript IV reverse transcriptase (Invitrogen) using an oligo(dT) primer. *CgACT2* gene was used as reference gene for normalization. qRT-PCR was performed on BIO-RAD CFX96 machine with gene-specific

forward and reverse primers (Additional file 1: Table S1). The reactions were carried out using Takara SYBR Premix Ex Taq II (Takara Bio Inc.) and incubated at 95 °C for 3 min followed by 40 cycles of 95 °C for 15 s, 58 °C for 15 s and 72 °C for 15 s. PCR specificity was checked by melting curve analysis, and data were analysed using the $2^{-\Delta\Delta CT}$ method [39].

Results

RNA sequencing, Trinity-based de novo transcriptome assembly and annotation using BLAST2GO and Trinotate

A total of 306,575,536 paired-end reads (150 bp) were obtained after sequencing all the twelve libraries on the Illumina NextSeq 500 platform. Subsequently, 186,399,131 good quality paired-end reads were used for de novo assembly of *Coccinia grandis* flower bud transcriptome using Trinity software package with default parameters (Table 1). The resulting assembly consisted of 467,233 'Trinity Transcripts' clustering into 378,860 'Trinity Genes' with an N50 value of 881 bp (Table 2, Additional file 2). The transcripts of 200–399 bp size were found to be most abundant in the length distribution of assembled transcripts (Fig. 2). However, a higher proportion of transcripts with length around 1000–2000 bp had a BLAST hit compared to the proportion of smaller transcripts (Fig. 2). Cleaned reads were mapped back to the transcriptome using bowtie2 with ~70% or more reads from each library aligning concordantly (Table 1). An Ex90N50 statistic calculated using 80,806 transcripts from the assembly (ignoring the rest of the transcripts with poor read coverage) was found to be 1784 bp (Additional file 3: Figure S1).

Altogether, 8916 unique BLAST hits in the Swiss-Prot database were represented by nearly full-length transcripts, having more than 70% alignment coverage, and 12,315 hits showed more than 50% alignment coverage (Additional file 4: Table S2). BUSCO output [C:89.8%(S:14.5%,D:75.3%),F:5.0%,M:5.2%,n:1440] showed that out of 1440 BUSCOs for Plants dataset, 1293 full length BUSCOs were detected in our de novo-assembled *Coccinia grandis* flower bud transcriptome indicating 89.8% completeness. Finally using TransRate, we were able to detect *C. grandis* homologs for 84% (18,039) of protein-coding primary transcripts of *C. sativus* of which, 13,430 reference sequences had at least 95% of their bases covered by a CRB-BLAST hit (Additional file 5: Table S3).

Coccinia grandis flower bud transcripts were compared to plant protein sequences of the nr and Swiss-Prot databases resulting in 259,200 and 136,663 transcripts having at least one hit from the respective database. Species distribution analysis of the BLAST hits showed that majority of these hits were from *Arabidopsis* and rice for Swiss-Prot database whereas for nr database most top hits were from cucumber and melon (Fig. 3a,b).

Devani *et al. BMC Plant Biology* (2017) 17:241

Page 5 of 15

Table 1 RNA sequencing read counts and alignment statistics for all the samples used for de novo transcriptome assembly

Sample name	Raw Reads	Cleaned reads	% Read pairs mapped concordantly
Male Early A	21,719,110	13,216,446	76.33%
Male Early B	22,080,977	12,149,633	74.84%
Female Early A	25,607,195	15,634,598	76.34%
Female Early B	25,433,955	15,316,996	75.94%
GyM-H Early A	27,936,206	17,378,706	76.67%
GyM-H Early B	27,617,808	17,107,330	76.69%
Ag-H Early A	26,147,527	15,140,736	71.02%
Ag-H Early B	25,392,540	15,128,539	69.35%
Male Middle A	25,502,209	15,490,027	76.70%
Male Middle B	25,837,400	15,919,727	75.27%
GyM-H Middle A	28,465,770	17,441,923	76.90%
GyM-H Middle B	24,834,839	16,474,470	77.43%
Total	306,575,536	186,399,131	

The number of transcripts annotated with various GO terms of biological process, molecular function, and cellular component categories are provided in Fig. 3c.

Trinotate v3 pipeline was also used simultaneously for comprehensive functional annotation of the *Coccinia grandis* flower bud transcripts. Details regarding the Swiss-Prot/TrEMBL BLAST hits, GO, KEGG and eggNOG mappings can be found in the Additional file 6: Table S4. HMMER/PFAM predicted protein domains, as well as information regarding signal peptides and transmembrane domains could also be found in the Additional file 6: Table S4. Taken together, we have assembled a good quality transcriptome for early- and middle-staged flower buds of *Coccinia grandis* and comprehensively annotated the transcripts using well-established BLAST2GO and Trinotate pipelines.

Differential expression analysis reveals probable factors for pollen fertility and sex modification

Following transcriptome assembly and annotation, differential expression analysis was carried out. First of all, RSEM was used for transcript abundance estimation.

Table 2 Assembly statistics for *C. grandis* flower bud transcriptome

Parameter	Assembly statistics
Number of 'Trinity Transcripts'	467,233
Number of 'Trinity Genes'	378,860
Percent GC	38.96
Median contig length (bp)	347
Average contig length (bp)	606.45
N50 (bp)	881
Total assembled bases	283,354,298

Following which, we checked for the correlation between the replicates for all the samples using *PtR* script. PCA analysis and correlation matrix showed a good correlation between the replicate sets for each of the six samples (Fig. 4). EdgeR was used to identify the differentially expressed transcripts for all the pairwise comparisons between the six samples (Table 3; Fig. 5; Additional file 7: Figure S2). Differentially expressed transcripts at a minimum fold change of 2^2 with p-values at most 1e-3 were extracted and GO enrichment analysis was performed (Additional file 8: Table S5, Additional file 9: Table S6). Among all the comparisons, few interesting ones such as Ag_Early_vs_Female_Early, Female_Early_vs_Male_Early, and GYM-H_Middle_vs_Male_Middle had 3574, 35,694 and 14,954 differentially expressed transcripts respectively (Table 3, Fig. 5). The DE features were partitioned into clusters with similar expression patterns (Fig. 6; Additional file 10: Figure S3). In the context of anther development, we identified several GO terms (GO:0080110, GO:0010208, GO:0010584, GO:0009555, GO:0055046, GO:0048658, GO:0048653) differentially enriched in male buds compared to female buds at an early stage of floral development. qRT-PCR was done to validate the results of differential expression analysis for a few interesting *Coccinia* homologs of *AMS* (ABORTED MICROSPORES), *CER3* (ECERIFERUM 3), *DEX1* (DEFECTIVE IN EXINE FORMATION 1), *DYT1* (DYSFUNCTIONAL TAPETUM 1), *EIL1* (ETHYLENE INSENSITIVE 3-like 1), *EMS1* (EXCESS MICROSPOROCYTES 1), *FER* (FERONIA), *MMD1* (MALE MEIOCYTE DEATH 1), *MS1* (MALE STERILITY 1), *SHT* (Spermidine hydroxycinnamoyl transferase), *TPD1* (TAPETUM DETERMINANT 1) and *ZAT3* (Zinc finger protein ZAT3). Expression profiles for these genes deduced by qRT-PCR revealed similar patterns to that seen from the digital DE analysis results (Fig. 7). Also, we have found GO terms related to pollen fertility enriched in the male buds (GO:0080092, GO:0009846, GO:0009860) compared to GyM-H and Ag-H buds, which had sterile pollens. Accordingly, expression profile for homologs of a number of genes involved in pollen tube development such as *CSLD1* (Cellulose synthase-like protein D1; TRINITY_DN92683_c0_g1_i1), *CDPKO* (Calcium-dependent protein kinase 24; TRINITY_DN93671_c0_g1_i3), *PME4* (Pectin methylesterase 4; TRINITY_DN14239_c0_g1_i1), *PME37* (Pectin methylesterase 37; TRINITY_DN3663_c0_g1_i1), *PPME1* (POLLEN SPECIFIC Pectin methylesterase 1; TRINITY_DN66415_c0_g1_i1) and *PTR52* (Protein NRT1/PTR FAMILY 2.8; TRINITY_DN112735_c0_g14_i3) were analysed and found to be enriched in middle-staged male buds similar to our digital expression profiles based on RNA-Seq data (Fig. 8). Downregulation of ethylene signalling upon AgNO$_3$ treatment was evident as GO:0009723 (response to ethylene) and

Devani et al. BMC Plant Biology (2017) 17:241

Page 6 of 15

Fig. 2 Transcript size distribution for *C. grandis* flower bud transcriptome
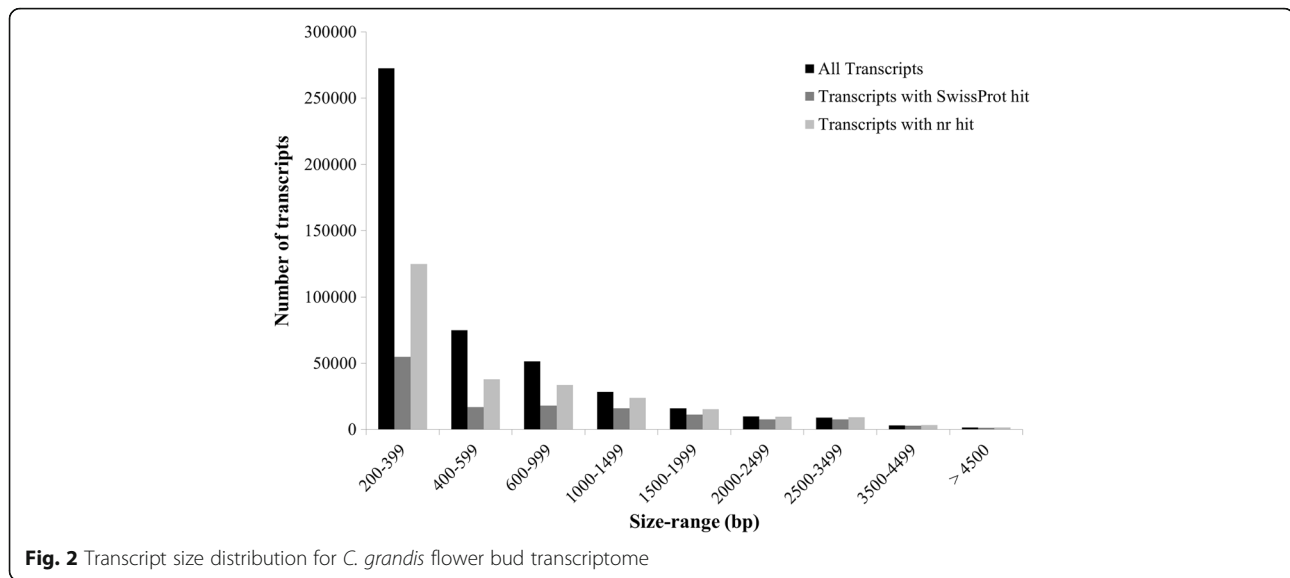
GO:0009873 (ethylene-activated signaling pathway) were depleted in AgNO$_3$ treated plant (Additional file 9: Table S6). In order to validate this, we studied the expression profile of *Ethylene-responsive transcription factors*, *ERF5*, *ERF17*, and *EF102*. We found that all three *ERFs* were downregulated upon AgNO$_3$ treatment (Fig. 9).

Discussion

Genetic basis of sex determination and differentiation is not well studied in *C. grandis*. Identification and investigation of sex-linked genes would lead to better understanding of dioecy in plants and this can be achieved by whole genome sequencing approach. However, sex-determining genes are most likely linked to non-recombining regions of Y-chromosome, which are difficult to assemble from sequence data [40]. An alternative approach is to use comparative transcriptomics to identify sex-biased genes that could play a role in sex differentiation and determination [24]. Further, the presence of mutations and SNPs in sex-biased genes can provide insights regarding the evolution of dioecy. Using this approach, we have assembled and annotated a de novo transcriptome from the flower buds of dioecious, gynomonoecious and AgNO$_3$ treated female *C. grandis*. We have identified differentially expressed genes which might be playing a role in stamen arrest of female flowers. Also, we have analysed the genes that were differentially expressed upon AgNO$_3$ treatment on female plants promoting stamen development. Finally, we have compared middle-staged male (bearing fertile pollens) and GyM-H buds (bearing sterile pollens) to study the genes involved in pollen maturation and fertility.

Differential expression of stamen developmental genes and arrest of stamen growth in female flowers

At the early stages (stages 3–4) of flower development in female *C. grandis*, both carpel and stamen organs are initiated simultaneously. However, stamen growth gets arrested during the course of development (stages 4–5) resulting in a female flower with rudimentary stamens. In contrast, no sign of carpel primordia was observed during the histological study of flower development in male *C. grandis* as described in our previous report [14]. The molecular players involved in stamen initiation and development process are well characterized in the hermaphrodite plant *Arabidopsis*. In order to identify the stage at which stamen growth gets arrested, *Coccinia grandis* homologs of *Arabidopsis* stamen development genes were identified from the de novo-assembled transcriptome. Among genes involved in stamen initiation, *Pistillata* (*CgPI*, TRINITY_DN71631_c0_g1_i1) was found to be expressed in a male-biased fashion (Additional file 8: Table S5). *Pistillata* has been shown to specify stamen identity in *Arabidopsis* [41] (Table 4). Further, EXCESS MICROSPOROCYTES 1 (EMS1) has been shown to interact with TAPETUM DETERMINANT 1 (TPD1) regulating specification of reproductive as well as somatic cells in *Arabidopsis* anthers [42]. Differential expression analyses revealed that homologs of both *EMS1* (TRINITY_DN106236_c0_g4_i1) and *TPD1* (TRINITY_DN116795_c2_g1_i3) were enriched in male flowers compared to female flowers (Table 4; Additional file 8: Table S5; Fig. 7). *DYSFUNCTIONAL TAPETUM 1* (*DYT1*) plays an important role in tapetum development by regulating the expression of *DEFECTIVE IN TAPETAL DEVELOPMENT AND FUNCTION 1* (*TDF1*) in *Arabidopsis* [43]. Also, DYT1 is known to interact with Basic helix-loop-

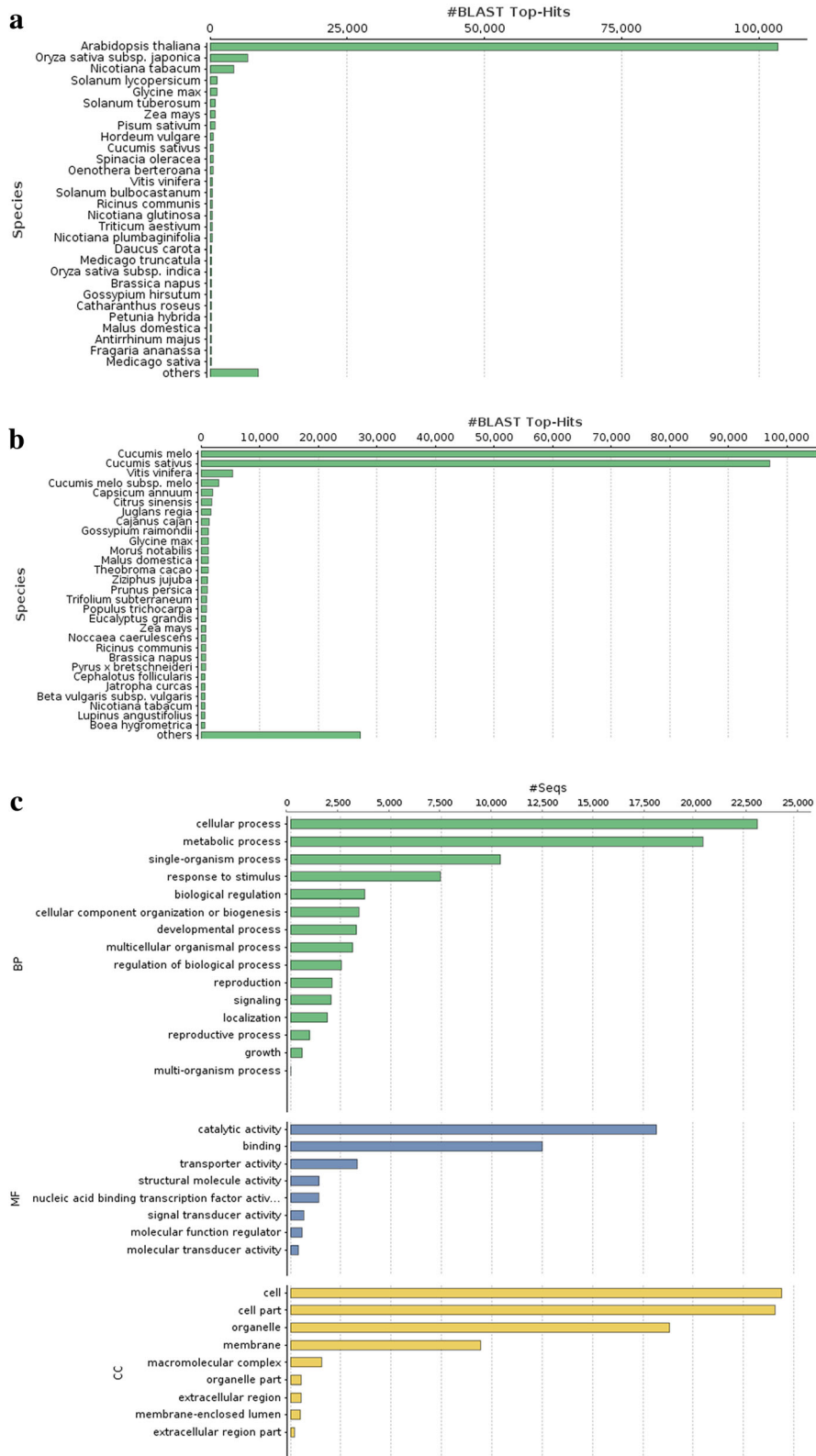Devani et al. BMC Plant Biology (2017) 17:241

Page 7 of 15

Fig. 3 BLAST2GO annotation of *C. grandis* flower bud transcriptome. (a) BLAST Top-Hits species distribution when compared with Swiss-Prot database, (b) BLAST Top-Hits species distribution when compared with nr database, (c) GO category distribution of *C. grandis* flower bud transcriptome

Fig. 4 Correlation analyses showing the relationship between samples and replicates. (**a**) Principal component analysis and (**b**) correlation matrix showing relationship between all samples as well as replicates
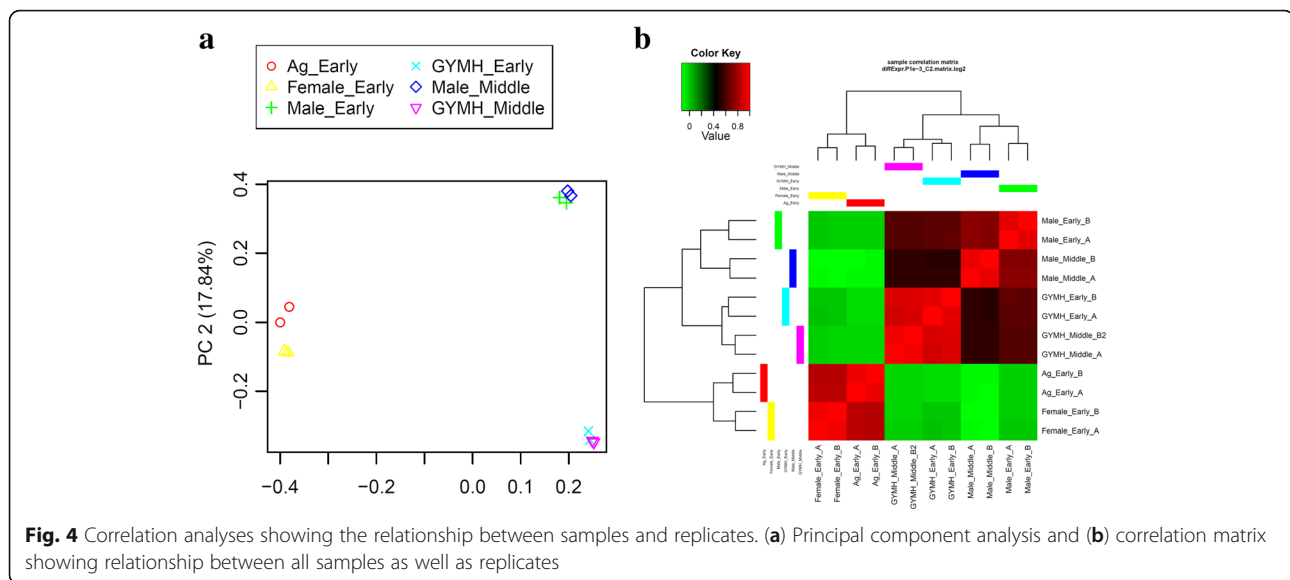
helix protein 89 (bHLH89) which is highly expressed in anthers and required for normal anther development and male fertility [44]. *TDF1* homolog (TRINITY_DN97604_c1_g7_i1) as well as *bHLH89* homolog (TRINITY_DN85771_c0_g1_i1) showed male-biased expression in *C. grandis* (Table 4; Additional file 8: Table S5). Differential regulation of these genes related to stamen development explains the possible cause for early stamen arrest in female flowers of *C. grandis*.

Table 3 Number of differentially expressed transcripts for each pairwise comparison between the flower types. Transcripts that had p-values at most 1e-3 and were at least 2^2 fold were considered as differentially expressed

Flower Buds Comparison	Number of DE transcripts
Ag_Early_vs_Female_Early	3574
Ag_Early_vs_GYMH_Early	34,458
Ag_Early_vs_GYMH_Middle	38,849
Ag_Early_vs_Male_Early	33,863
Ag_Early_vs_Male_Middle	36,923
Female_Early_vs_GYMH_Early	31,886
Female_Early_vs_GYMH_Middle	39,885
Female_Early_vs_Male_Early	35,694
Female_Early_vs_Male_Middle	40,477
GYMH_Early_vs_GYMH_Middle	816
GYMH_Early_vs_Male_Early	8659
GYMH_Early_vs_Male_Middle	11,576
GYMH_Middle_vs_Male_Early	12,357
GYMH_Middle_vs_Male_Middle	14,954
Male_Early_vs_Male_Middle	4427

According to recent reports from monoecious cucurbits like melon, cucumber, and watermelon, ethylene plays a major role in sex determination by inhibiting stamen development process [45–48]. We found that compared to male, GO:0009723 (response to ethylene) was enriched in female buds indicating a potential role of ethylene in sex determination of *C. grandis* (Additional file 9: Table S6).

AgNO₃ treatment on female plant releases the stamen inhibition

Female plants of *C. grandis* bear flowers with fused carpels and rudimentary stamens. Earlier, we have shown that foliar spray of 35 mM AgNO₃ on the female plant of *C. grandis* promotes further development of the rudimentary stamens [14]. In the current study, gene expression profiles for early-staged Ag-H flower buds were compared with female buds (Table 3; Fig. 5c, g). Ag⁺ ions are known to inhibit responses to ethylene, a gaseous plant hormone [17]. Also, silver compounds have been shown to induce maleness by promoting stamen development in many monoecious and dioecious species [19–21]. No other inhibitors of ethylene biosynthesis or signaling could induce the stamen development in *Silene latifolia*, suggesting that ethylene signaling might not be the only pathway that gets affected upon application of silver thiosulphate [21]. In contrast to *Silene latifolia*, AVG (aminoethoxyvinylglycine), an inhibitor of ethylene-biosynthesis has been shown to induce male flowers in gynoecious muskmelon similar to silver compounds [49]. Considering the role of 1-aminocyclopropane-1-carboxylate synthase (ACS, an enzyme involved in ethylene biosynthesis) in sex determination of many other members of Cucurbitaceae, an ethylene-mediated effect of

Devani et al. BMC Plant Biology (2017) 17:241

Page 9 of 15

Fig. 5 Pairwise comparisons of transcript abundance. MA plots showing average log fold change (logFC) vs average log of counts among (**a**) female (early-staged) vs. male (early-staged) transcripts, (**b**) female (early-staged) vs. GyM-H (early-staged) transcripts, (**c**) Ag-H (early-staged) vs. female (early-staged) transcripts and (**d**) GyM-H (middle-staged) vs. male (middle-staged) across replicates. Volcano plots showing differentially expressed transcripts in relation to FDR (False discovery rate) for (**e**) female (early-staged) vs. male (early-staged) transcripts, (**f**) female (early-staged) vs. GyM-H (early-staged) transcripts, (**g**) Ag-H (early-staged) vs. female (early-staged) transcripts and (**h**) GyM-H (middle-staged) vs. male (middle-staged). Features found DE at FDR <0.05 are colored red. Features with P-values at most 1e-3 and at least 2^2 fold change are differentially expressed

$AgNO_3$ seems more likely to be involved in the modification of sex in *C. grandis* [50].

In our study, GO:0009723 (response to ethylene) and GO:0009873 (ethylene-activated signaling pathway) were enriched in female buds compared to Ag-H buds (Additional file 9: Table S6). Transcripts for genes such as *Ethylene-responsive transcription factors, ERF5* (TRINITY_DN102355_c3_g13_i1), *ERF17* (TRINITY_DN80749_c0_g6_i1), EF109 (TRINITY_DN87049_c0_g1_i1), EF102 (TRINITY_DN90257_c1_g2_i1), ERF99 (TRINITY_DN93821_c0_g1_i2), ERF60 (TRINITY_DN93262_c1_g6_i2) and *ERF78 (TRINITY_DN98503_c3_g1_i1)* were downregulated in Ag-H buds indicating impaired ethylene signalling (Additional file 8: Table S5). Additionally, qRT-PCR based expression pattern analysis for *ERF5, ERF17* and *EF102* genes clearly showed the suppression of ethylene responses by $AgNO_3$ (Fig. 9).

Downregulation of ethylene signaling in Ag-H buds was correlated with the promotion of stamen growth. GO:0048655 (anther wall tapetum morphogenesis), GO:0048657 (anther wall tapetum cell differentiation), GO:0048658 (anther wall tapetum development) were seen to be enriched in early-staged Ag-H buds compared to female buds (Additional file 9: Table S6). *C. grandis* homologs of *MS1, MMD1* (TRINITY_DN109512_c4_g3_i1, TRINITY_DN108927_c0_g6_i1), *ZAT3* (TRINITY_DN10

8658_c0_g2_i1) and *AMS* (TRINITY_DN116105_c0_g2_i1) genes which play important roles in tapetum and pollen development of *Arabidopsis* flowers were upregulated upon $AgNO_3$ treatment indicating promotion of stamen growth [51–56] (Additional file 8: Table S5; Fig. 7). *MYB35* (TRINITY_DN92649_c0_g7_i1), which was proposed as a putative sex-determining gene in *Asparagus* was also found to be upregulated in Ag-H buds [57] (Additional file 8: Table S5). Apart from that, gene ontology terms related to pollen wall assembly (GO:0010208), pollen exine formation (GO:0010584), sporopollenin biosynthetic process (GO:0080110), pollen development (GO:0009555) and pollen sperm cell differentiation (GO:0048235) were also enriched in Ag-H buds (Additional file 9: Table S6). Further, we noticed that *Ethylene-responsive transcription factors* (ERFs) were not affected in GyM-H buds as compared to female buds suggesting that stamen development in GyM-H flower buds might be regulated by some other mechanism evading ethylene signaling inhibition.

Transcripts governing pollen fertility are depleted in GyM-H and Ag-H flower buds

C. grandis is one of the few species in which the presence of heteromorphic sex chromosomes is reported. The large Y-chromosome present in males might play a major role in sex determination. The GyM form of *C.*

Devani et al. BMC Plant Biology (2017) 17:241

Page 10 of 15
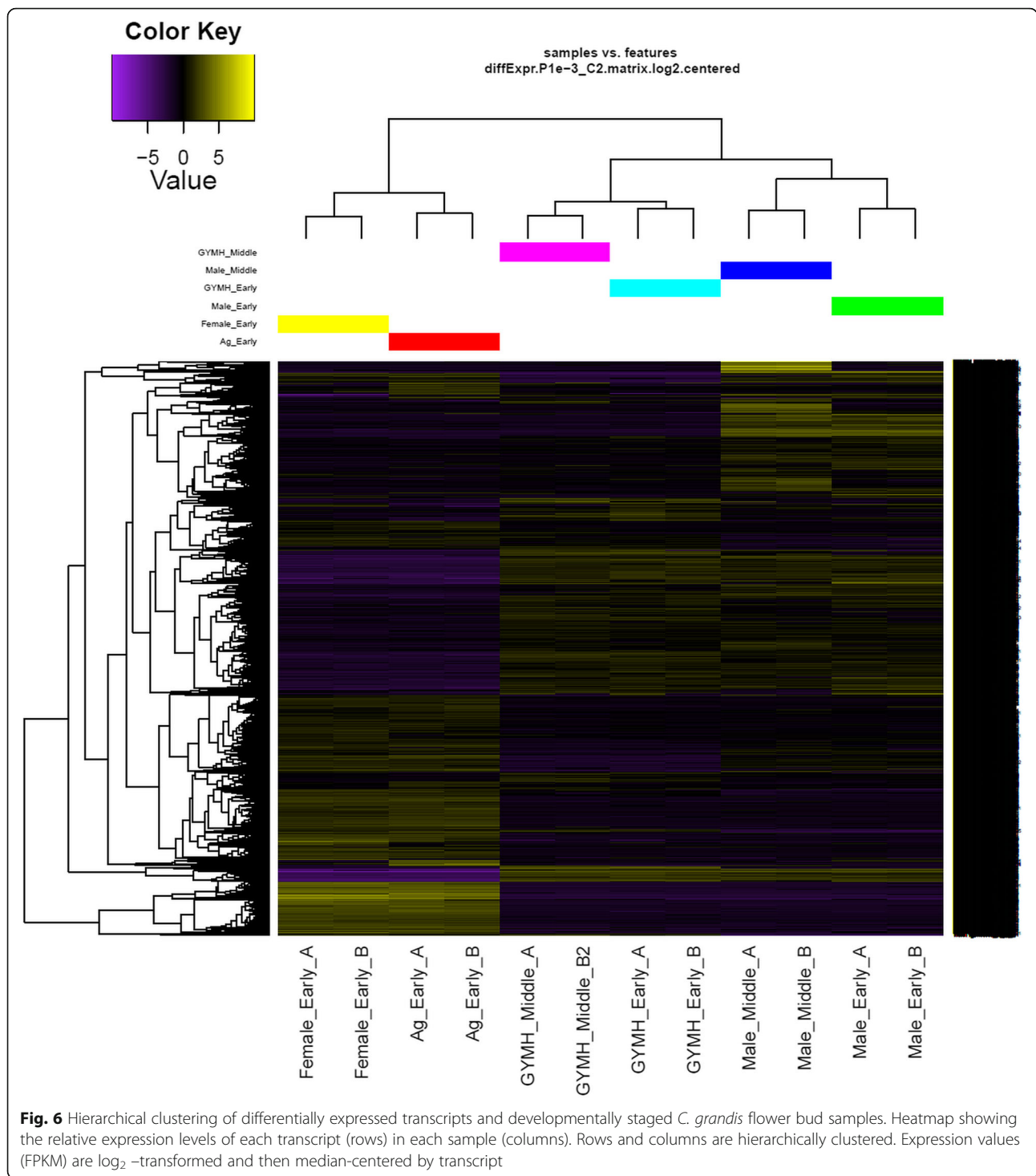
Fig. 6 Hierarchical clustering of differentially expressed transcripts and developmentally staged *C. grandis* flower bud samples. Heatmap showing the relative expression levels of each transcript (rows) in each sample (columns). Rows and columns are hierarchically clustered. Expression values (FPKM) are log$_2$ –transformed and then median-centered by transcript

grandis included in the current study does not have Y-chromosome [14]. GyM-H flowers still develop full-sized stamens despite lacking Y-chromosome. Similarly, AgNO$_3$ treatment induces stamen development in female plants having XX sex chromosomes. However, the pollens from GyM-H and Ag-H flowers buds were found to be sterile unlike the pollens from male buds [14].

Differential expression analysis revealed that gene ontology terms for pollen tube (GO:0090406), pollen germination (GO:0009846), regulation of pollen tube growth (GO:0080092), pollen tube growth (GO:0009860) and microsporogenesis (GO:0009556) were enriched in middle-staged male buds compared to middle-staged GyM-H buds (Additional file 9: Table S6).

Devani et al. BMC Plant Biology (2017) 17:241

Page 11 of 15

Fig. 7 Validation of selected DE genes by qRT-PCR with two biological replicates. The relative expression in the sample of middle-staged male-A (MMA) was set to 1 for plotting the qRT-PCR data. AgEA, Early-staged Ag-H A; AgEB, Early-staged Ag-H B; FEA, Early-staged Female A; FEB, Early-staged Female B; GEA, Early-staged GyM-H A; GEB, Early-staged GyM-H B; GMA, Middle-staged GyM-H A; GMB, Middle-staged GyM-H B; MEA, Early-staged Male A; MEB, Early-staged Male B; MMA, Middle-staged Male A; MMB, Middle-staged Male B; AgMA, Middle-staged Ag-H A, AgMB, Middle-staged Ag-H B; FMA, Middle-staged Female A; FMB, Middle-staged Female B
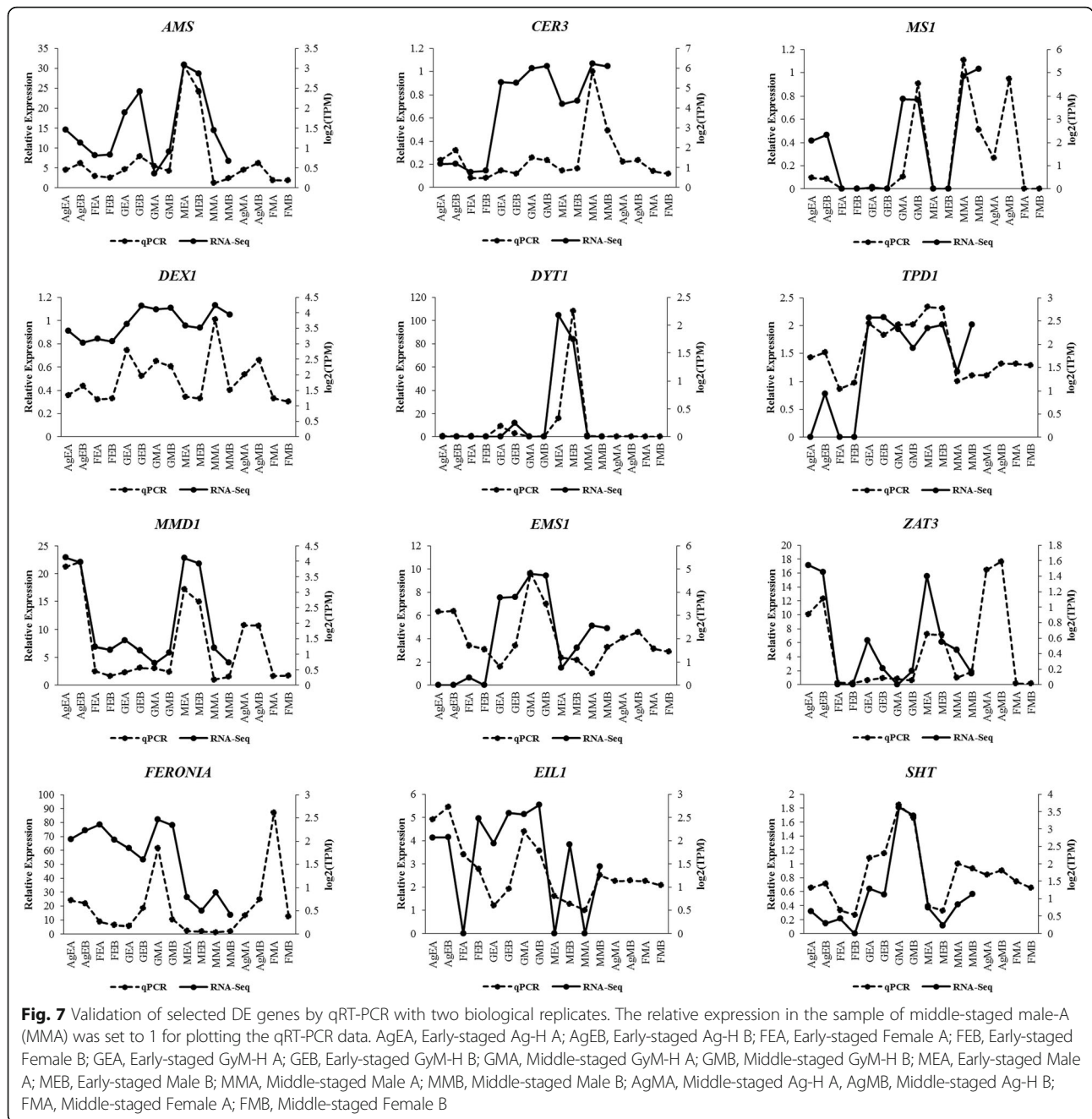
GAUTE plays an important role in pollen tube wall biosynthesis in *Arabidopsis* [58]. TRINITY_DN111340_c1_g1_i6, which showed similarity with *GAUTE* was enriched in male buds compared to GyM-H buds. Unlike most other plant cell walls, pollen tube wall does not contain callose or cellulose. Pectin methylesterases (PMEs) have been shown to play a very important role in the growth of pollen tubes [59–61]. *PME4* (TRINITY_DN14239_c0 _g1_i1), *PME37* (TRINITY_DN3663_c0_g1_i1) and *PPME1* (TRINITY_DN66415_c0_g1_i1, TRINITY_DN7 1598_c0_g2_i1) were downregulated in GyM-H buds

compared to male buds (Additional file 8: Table S5). This could be a possible cause for pollens from GyM-H not forming pollen tubes. H Zhan, Y Zhong, Z Yang and H Xia [62] has shown that *IPMKB* (Inositol polyphosphate multikinase beta) is an important factor for pollen development. We have found that TRINITY_DN96290_c0_g3_i2 transcript matching to Arabidopsis *IPMKB* (*AtIpk2beta*) was downregulated in GyM-H compared to male buds. Earlier, several reports have demonstrated that *MALE STERILITY 1* (*MS1*) gene of *Arabidopsis* expresses in tapetal cells and plays

Devani *et al. BMC Plant Biology* (2017) 17:241

Page 12 of 15

Fig. 8 qRT-PCR based expression analyses of selected genes involved in pollen tube development with two biological replicates. The relative expression in the sample of middle-staged male-A (MMA) was set to 1 for plotting the qRT-PCR data. AgEA, Early-staged Ag-H A; AgEB, Early-staged Ag-H B; FEA, Early-staged Female A; FEB, Early-staged Female B; GEA, Early-staged GyM-H A; GEB, Early-staged GyM-H B; GMA, Middle-staged GyM-H A; GMB, Middle-staged GyM-H B; MEA, Early-staged Male A; MEB, Early-staged Male B; MMA, Middle-staged Male A; MMB, Middle-staged Male B; AgMA, Middle-staged Ag-H A, AgMB, Middle-staged Ag-H B; FMA, Middle-staged Female A; FMB, Middle-staged Female B
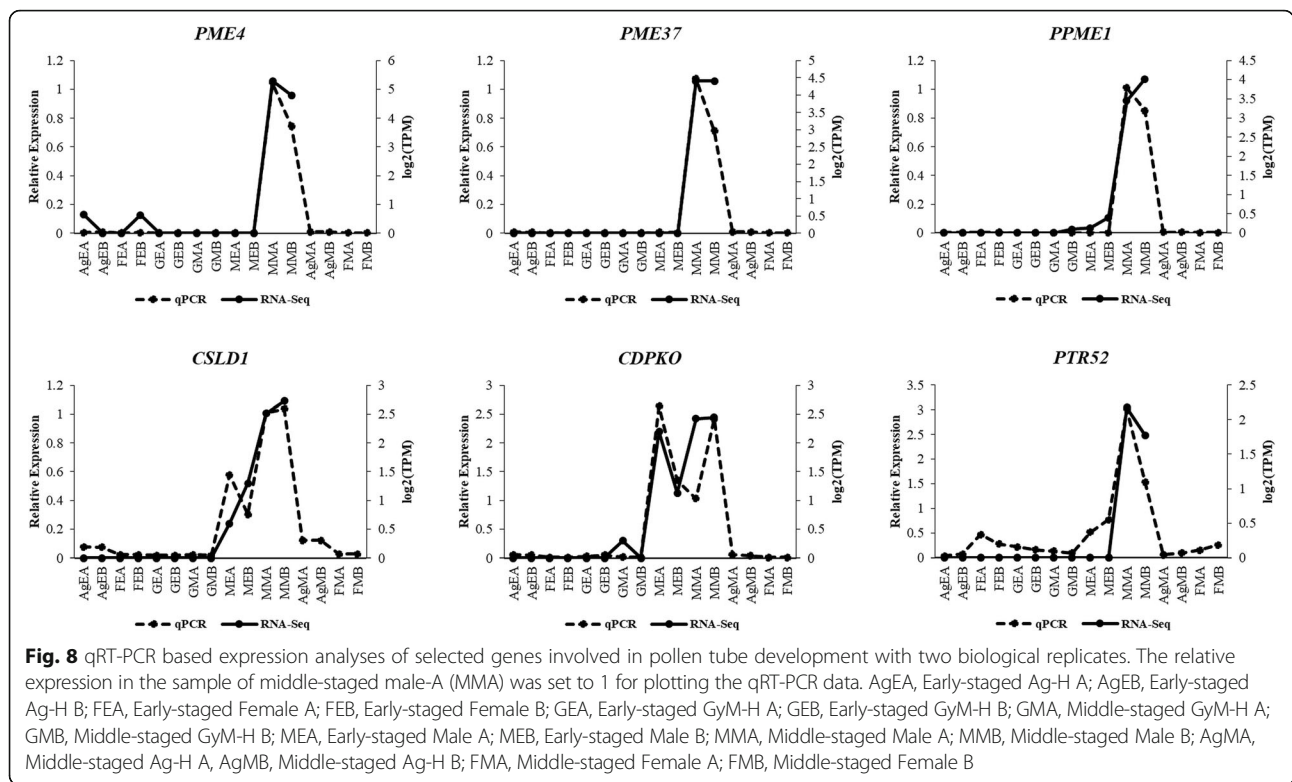
an important role in pollen maturation [51, 54, 55]. *C. grandis* homolog of *MS1*, TRINITY_DN109512_c4 _g3_i1 was expressed in a male-biased manner (Additional file 8: Table S5; Fig. 7). Similarly, homologs of genes important for pollen tube growth such as *CSLD1* (TRINITY_DN92683_c0_g1_i1), *CDPKO* (TRINITY_DN93671_c0_g1_i3), *NRX1* (TRINITY_DN 106708_c1_g2_i3), *PTR52* (TRINITY_DN112735_c0_g 14_i3; TRINITY_DN112735_c0_g3_i1), *TAF6* (TRINI TY_DN96231_c1_g1_i2) and *CALS5* (TRINITY_DN1 13564_c1_g1_i1) were enriched in male [63–68] (Fig. 8, Additional file 8: Table S5). Genes involved

in pollen exine formation such as *FACR2/MS2* (TRI NITY_DN74585_c1_g5_i3), *EA6* (TRINITY_DN7627 4_c1_g1_i1), *C70A2/DEX2* (TRINITY_DN99059_c0_ g1_i1) were also upregulated in male [69–71]. *EMS1* (TRINITY_DN89942_c0_g7_i1), *SERK1* (TRINITY_D N108624_c1_g7_i5), *JASON* (TRINITY_DN83440_c0 _g1_i1), *RPK2* (TRINITY_DN113423_c0_g1_i4), which are essential for microsporogenesis and pollen matur-ation were observed to be expressed at significantly higher levels in middle-staged male buds compared to GyM-H buds. [72–75] (Additional file 8: Table S5; Fig. 7).

Fig. 9 qRT-PCR based expression analyses of selected *Ethylene-responsive transcription factors* (*ERFs*) with two biological replicates. The relative expression in the sample of middle-staged male-A (MMA) was set to 1 for plotting the qRT-PCR data. AgEA, Early-staged Ag-H A; AgEB, Early-staged Ag-H B; FEA, Early-staged Female A; FEB, Early-staged Female B; GEA, Early-staged GyM-H A; GEB, Early-staged GyM-H B; GMA, Middle-staged GyM-H A; GMB, Middle-staged GyM-H B; MEA, Early-staged Male A; MEB, Early-staged Male B; MMA, Middle-staged Male A; MMB, Middle-staged Male B; AgMA, Middle-staged Ag-H A, AgMB, Middle-staged Ag-H B; FMA, Middle-staged Female A; FMB, Middle-staged Female B

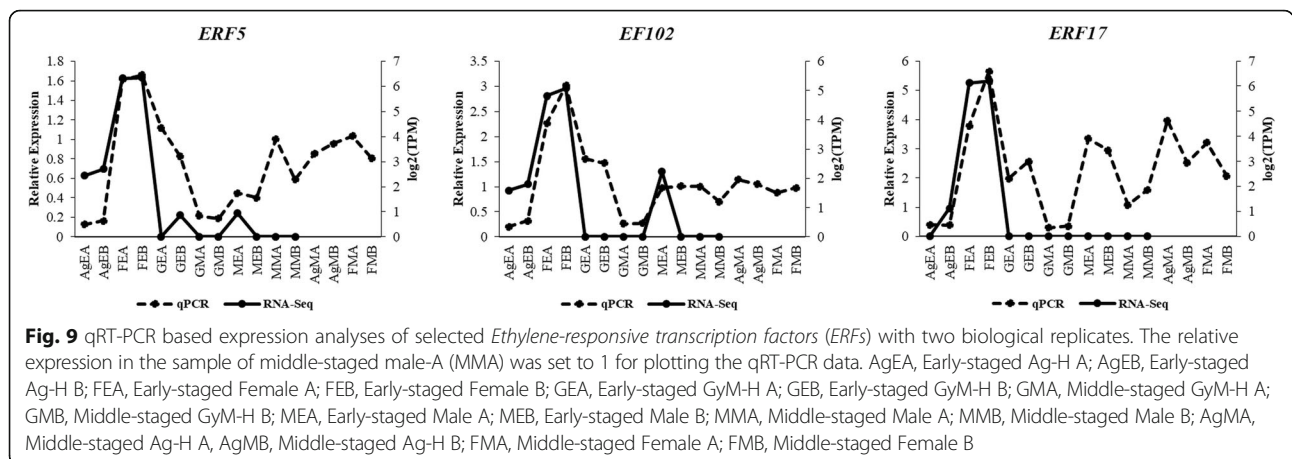Devani *et al. BMC Plant Biology* (2017) 17:241

Page 13 of 15

Table 4 Digital Expression profile for genes in anther developmental pathway

Stage of stamen development	Genes	Expression pattern
Stamen Primordia Initiation	AG	Unbiased
	CLV1/CLV2	Unbiased/male-biased
	PI	Male-biased
	AP3	Unclear homolog
	JAG	Male-biased
Archesporial initiation	BAM1/BAM2	Unclear homolog
	SPL/NZZ	Unclear homolog
Tapetal Development	EMS1	Male-biased
	SERK1/2	Male-specific
	TPD1	Male-biased
	RPK2	Male-specific
	TDF	Male-biased
	DYT1	Unbiased
	bHLH89	Male-biased
Mature Pollen Formation	AMS	Male-biased
	MS1	Male-biased
	MS2	Male-biased
	MIA	Unbiased
	LAP3	Unbiased
	LAP5	Male-biased

Expression profiling for homologs of *MS1, EMS1, DYT1, PME4, PME37, PPME1, CSLD1, CDPKO,* and *PTR52* was studied by qRT-PCR for all the tissue samples including middle-staged Ag-H buds (Figs. 7 and 8; Table 4). Transcripts for all these homologs were downregulated in Ag-H buds and GyM-H buds, suggesting a male-biased expression pattern implicating the reason for pollen sterility in Ag-H and GyM-H buds.

Conclusions

De novo-assembled transcriptome developed from RNA-Seq of different sexual phenotypes has enabled identification of *C. grandis* homologs of many genes known to be involved in flower development in species such as *Arabidopsis*, melon, and cucumber. We found out that many genes involved in stamen initiation, tapetal development, and pollen maturation were downregulated in female buds compared to male buds. Interestingly, ethylene response-related genes were upregulated in female buds compared to male buds indicating a probable role of ethylene in stamen suppression similar to monoecious cucurbits such as melon and cucumber. We speculate that the Y-chromosome might express genes that inhibit ethylene signaling or suppress the carpel development, the site

of ethylene production leading to the formation of stamens in male flowers. This was supported by the observation that AgNO$_3$ treatment suppressed ethylene responses and induced stamen development in female flowers of *C. grandis*. However, the pollens produced by Ag-H flowers were sterile indicating a decisive role of Y-chromosome in determining maleness. In accordance with this, the transcripts involved in pollen maturation, pollen germination, and pollen tube elongation were depleted in middle-staged GyM-H buds compared to male buds. This could be because of the absence of Y-chromosome in GyM plant. Altogether, differentially expressed genes identified in this study could shed light on the probable mechanisms of sex determination, differentiation, and modification in *C. grandis*.

Additional files

Abbreviations

Devani *et al. BMC Plant Biology* (2017) 17:241

Page 14 of 15

PCR: Quantitative real-time PCR; RIN: RNA integrity number; RSEM: RNA-Seq by Expectation-Maximization; Swiss-Prot: Annotated protein sequence database; TrEMBL: Computer-annotated supplement to Swiss-Prot database

Availability of data and materials
The raw sequencing data were deposited in the NCBI Short Read Archive (SRA) database (http://www.ncbi.nlm.nih.gov/sra/) under the accession number SRP111347. De novo-assembled *Coccinia grandis* floral bud transcriptome is available at https://doi.org/10.6084/m9.figshare.5220874.v1. Annotation Report for the de novo-assembled *C. grandis* flower bud transcriptome generated using Trinotate is available at https://doi.org/10.6084/m9.figshare.5217652.v1.

Author Contributions
RSD, AKB, and SS planned and designed the research, RSD performed experiments and analysed the data. RSD, AKB, and SS wrote the manuscript. AB, JB, and RKS helped in critical discussion, interpretation of data and edited the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Biology Division, Indian Institute of Science Education and Research (IISER), Pune, Pune, Maharashtra, India. [2]Department of Botany, Tripura University, Suryamaninagar, Tripura, India. [3]IPS2, INRA, CNRS, University Paris Sud, University of Evry, University Paris Diderot, University of Paris Saclay, Batiment 630, 91405 Orsay, France.

References
1. Charlesworth D. Plant sex determination and sex chromosomes. Heredity. 2002;88(2):94–101.
2. Ainsworth C, Crossley S, Buchanan-Wollaston V, Thangavelu M, Parker J. Male and female flowers of the dioecious plant sorrel show different patterns of MADS box gene expression. Plant Cell. 1995;7(10):1583–98.
3. Pfent C, Pobursky KJ, Sather DN, Golenberg EM. Characterization of SpAPETALA3 and SpPISTILLATA, B class floral identity genes in Spinacia oleracea, and their relationship to sexual dimorphism. Dev Genes Evol. 2005;215(3):132–42.
4. Yin TM, DiFazio SP, Gunter LE, Zhang X, Sewell MM, Woolbright SA, Allan GJ, Kelleher CT, Douglas CJ, Wang M, et al. Genome structure and emerging evidence of an incipient sex chromosome in Populus. Genome Res. 2008;18(3):422–30.
5. Urasaki N, Tarora K, Shudo A, Ueno H, Tamaki M, Miyagi N, Adaniya S, Matsumura H. Digital transcriptome analysis of putative sex-determination genes in papaya (Carica papaya). PLoS One. 2012;7(7):e40904.
6. Kouonon LC, Jacquemart A-L, Zoro Bi AI, Bertin P, Baudoin J-P, Dje Y. Reproductive biology of the andromonoecious Cucumis melo subsp. agrestis (Cucurbitaceae). Ann Bot. 2009;104(6):1129–39.
7. Kumar LSS, Deodikar GB. Sex chromosomes of Coccinia indica Wight and Arn. Curr Sci. 1940;9(3):128–30.
8. Bhaduri PN, Bose PC. Cyto-genetical investigations in some common cucurbits, with special reference to fragmentation of chromosomes as a physical basis of speciation. J Genet. 1947;48(2):237–56.
9. Chakravorti AK. Cytology of Coccinia indica W. & a. With reference to the behaviour of its sex-chromosomes. Proceedings: Plant Sciences. 1948;27(3):74–86.
10. Kumar LS, Viseveshwaraiah S. Sex mechanism in Coccinia indica Wight and Arn. Nature. 1952;170(4321):330–1.
11. Bhowmick BK, Jha TB, Jha S. Chromosome analysis in the dioecious cucurbit Coccinia grandis (L.) Voigt. Chromosome Science. 2012;15(1+ 2):9–15.
12. Sousa A, Fuchs J, Renner SS. Molecular cytogenetics (FISH, GISH) of Coccinia grandis: a ca. 3 myr-old species of Cucurbitaceae with the largest Y/autosome divergence in flowering plants. Cytogenetic and genome research. 2013;139(2):107–18.
13. Bhuskute S, Makde K, Deshpande P. Staminal organization in Coccinia grandis (L.) Voigt. Ann Bot. 1986:415–8.
14. Ghadge AG, Karmakar K, Devani RS, Banerjee J, Mohanasundaram B, Sinha RK, Sinha S, Banerjee AK. Flower development, pollen fertility and sex expression analyses of three sexual phenotypes of Coccinia grandis. BMC Plant Biol. 2014;14(1):325.
15. Hardenack S, Ye D, Saedler H, Grant S. Comparison of MADS box gene expression in developing male and female flowers of the dioecious plant white campion. Plant Cell. 1994;6(12):1775–87.
16. Di Stilio VS, Kramer EM, Baum DA. Floral MADS box genes and homeotic gender dimorphism in Thalictrum dioicum (Ranunculaceae) – a new model for the study of dioecy. Plant J. 2005;41(5):755–66.
17. Beyer EM. A potent inhibitor of ethylene action in plants. Plant Physiol. 1976;58(3):268–71.
18. McDaniel BK, Binder BM. Ethylene receptor 1 (ETR1) is sufficient and has the predominant role in mediating inhibition of ethylene responses by silver in Arabidopsis thaliana. J Biol Chem. 2012;287(31):26094–103.
19. Sarath G, Mohan Ram HY. Comparative effect of silver ion and gibberellic acid on the induction of male flowers on female Cannabis plants. Experientia. 1979;35(3):333–4.
20. Yin T, Quinn JA. Tests of a mechanistic model of one hormone regulating both sexes in Cucumis sativus (Cucurbitaceae). Am J Bot. 1995:1537–46.
21. Law TF, Lebel-Hardenack S, Grant SR. Silver enhances stamen development in female white campion (Silene latifolia [Caryophyllaceae]). Am J Bot. 2002;89(6):1014–20.
22. Blackburn KB. Sex chromosomes in plants. Nature. 1923;112(2819):687–8.
23. Kihara H, Ono T. The sex-chromosomes of Rumex acetosa. Zeitschrift für induktive Abstammungs-und Vererbungslehre. 1925;39(1):1–7.
24. Muyle A, Zemp N, Deschamps C, Mousset S, Widmer A, Marais GA. Rapid de novo evolution of X chromosome dosage compensation in Silene latifolia, a plant with young sex chromosomes. PLoS Biol. 2012;10(4):e1001308.
25. Wu T, Qin Z, Zhou X, Feng Z, Du Y. Transcriptome profile analysis of floral sex determination in cucumber. J Plant Physiol. 2010;167(11):905–13.
26. Guo S, Zheng Y, Joung J-G, Liu S, Zhang Z, Crasta OR, Sobral BW, Xu Y, Huang S, Fei Z. Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. BMC Genomics. 2010;11(1):384.
27. Akagi T, Henry IM, Tao R, Comai L. A Y-Chromosome–encoded small RNA acts as a sex determinant in persimmons. Science. 2014;346(6209):646–50.
28. Harkness A, Mercati F, Shan HY, Sunseri F, Falavigna A, Leebens-Mack J. Sex-biased gene expression in dioecious garden asparagus (Asparagus officinalis). New Phytol. 2015;207(3):883–92.
29. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014:btu170.
30. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol. 2011;29(7):644.
31. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. De novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. Nat Protoc. 2013;8(8):1494–512.
32. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.

Devani *et al. BMC Plant Biology* (2017) 17:241

Page 15 of 15

33. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015:btv351.

34. Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S. TransRate: reference-free quality assessment of *de novo* transcriptome assemblies. Genome Res. 2016;26(8):1134–44.

35. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

36. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 2008;36(10):3420–35.

37. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics. 2011;12(1):323.

38. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.

39. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 2001;25(4): 402–8.

40. Muyle A, Käfer J, Zemp N, Mousset S, Picard F, Marais GA. SEX-DETector: a probabilistic approach to study sex chromosomes in non-model organisms. Genome Biology and Evolution. 2016;8(8):2530–43.

41. Krizek BA, Meyerowitz EM. The *Arabidopsis* homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. Development. 1996;122(1):11–22.

42. Jia G, Liu X, Owen HA, Zhao D. Signaling of cell fate determination by the TPD1 small protein and EMS1 receptor kinase. Proc Natl Acad Sci. 2008; 105(6):2220–5.

43. Gu JN, Zhu J, Yu Y, Teng XD, Lou Y, Xu XF, Liu JL, Yang ZN. DYT1 directly regulates the expression of *TDF1* for tapetum development and pollen wall formation in *Arabidopsis*. Plant J. 2014;80(6):1005–13.

44. Zhu E, You C, Wang S, Cui J, Niu B, Wang Y, Qi J, Ma H, Chang F. The DYT1-interacting proteins bHLH010, bHLH089 and bHLH091 are redundantly required for *Arabidopsis* anther development and transcriptome. Plant J. 2015;83(6):976–90.

45. Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, Sari M-A, Collin F, Flowers JM, Pitrat M, Purugganan MD, Dogimont C, Bendahmane A. A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science. 2008;321(5890):836–8.

46. Boualem A, Troadec C, Kovalski I, Sari M-A, Perl-Treves R, Bendahmane A. A conserved ethylene biosynthesis enzyme leads to andromonoecy in two *Cucumis* species. PLoS One. 2009;4(7):e6144.

47. Boualem A, Troadec C, Camps C, Lemhemdi A, Morin H, Sari M-A, Fraenkel-Zagouri R, Kovalski I, Dogimont C, Perl-Treves R, Bendahmane A. A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. Science. 2015;350(6261):688–91.

48. Boualem A, Lemhemdi A, Sari M-A, Pignoly S, Troadec C, Abou Choucha F, Solmaz I, Sari N, Dogimont C, Bendahmane A. The andromonoecious sex determination gene predates the separation of *Cucumis* and *Citrullus* genera. PLoS One. 2016;11(5):e0155444.

49. Owens K, Peterson C, Tolla G. Production of hermaphrodite flowers on gynoecious muskmelon by silver nitrate and aminoethyoxyvinylglycine. Hortscience. 1980;15(5):654–5.

50. Zhang J, Boualem A, Bendahmane A, Ming R. Genomics of sex determination. Curr Opin Plant Biol. 2014;18:110–6.

51. Ito T, Shinozaki K. The *MALE STERILITY1* gene of *Arabidopsis*, encoding a nuclear protein with a PHD-finger motif, is expressed in tapetal cells and is required for pollen maturation. Plant Cell Physiol. 2002;43(11):1285–92.

52. Sorensen AM, Kröber S, Unte US, Huijser P, Dekker K, Saedler H. The *Arabidopsis ABORTED MICROSPORES* (*AMS*) gene encodes a MYC class transcription factor. Plant J. 2003;33(2):413–23.

53. Yang X, Makaroff CA, Ma H. The *Arabidopsis MALE MEIOCYTE DEATH1* gene encodes a PHD-finger protein that is required for male meiosis. Plant Cell. 2003;15(6):1281–95.

54. Ito T, Nagata N, Yoshiba Y, Ohme-Takagi M, Ma H, Shinozaki K. *Arabidopsis MALE STERILITY1* encodes a PHD-type transcription factor and regulates pollen and tapetum development. Plant Cell. 2007;19(11):3549–62.

55. Yang C, Vizcay-Barrena G, Conner K, Wilson ZA. MALE STERILITY1 is required for tapetal development and pollen wall biosynthesis. Plant Cell. 2007; 19(11):3530–48.

56. Borg M, Brownfield L, Khatab H, Sidorova A, Lingaya M, Twell D. The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in *Arabidopsis*. Plant Cell. 2011;23(2): 534–49.

57. Tsugama D, Matsuyama K, Ide M, Hayashi M, Fujino K, Masuda K. A putative MYB35 ortholog is a candidate for the sex-determining genes in *Asparagus officinalis*. Sci Rep. 2017;7

58. Wang L, Wang W, Wang Y-Q, Liu Y-Y, Wang J-X, Zhang X-Q, Ye D, Chen L-Q. *Arabidopsis* galacturonosyltransferase (GAUT) 13 and GAUT14 have redundant functions in pollen tube growth. Mol Plant. 2013;6(4):1131–48.

59. Jiang L, Yang S-L, Xie L-F, Puah CS, Zhang X-Q, Yang W-C, Sundaresan V, Ye D. *VANGUARD1* encodes a pectin methylesterase that enhances pollen tube growth in the *Arabidopsis* style and transmitting tract. Plant Cell. 2005;17(2): 584–96.

60. Tian G-W, Chen M-H, Zaltsman A, Citovsky V. Pollen-specific pectin methylesterase involved in pollen tube growth. Dev Biol. 2006;294(1):83–91.

61. Leroux C, Bouton S, Kiefer-Meyer M-C, Fabrice TN, Mareck A, Guénin S, Fournet F, Ringli C, Pelloux J, Driouich A, Lerouge P, Lehner A, Mollet J-C. PECTIN METHYLESTERASE48 is involved in *Arabidopsis* pollen grain germination. Plant Physiol. 2015;167(2):367–80.

62. Zhan H, Zhong Y, Yang Z, Xia H. Enzyme activities of *Arabidopsis* inositol polyphosphate kinases AtIPK2α and AtIPK2β are involved in pollen development, pollen tube guidance and embryogenesis. Plant J. 2015;82(5):758–71.

63. Dong X, Hong Z, Sivaramakrishnan M, Mahfouz M, Verma DPS. Callose synthase (CalS5) is required for exine formation during microgametogenesis and for pollen viability in *Arabidopsis*. Plant J. 2005;42(3):315–28.

64. Lago C, Clerici E, Dreni L, Horlow C, Caporali E, Colombo L, Kater MM. The *Arabidopsis* TFIID factor AtTAF6 controls pollen tube growth. Dev Biol. 2005; 285(1):91–100.

65. Bernal AJ, Yoo C-M, Mutwil M, Jensen JK, Hou G, Blaukopf C, Sørensen I, Blancaflor EB, Scheller HV, Willats WG. Functional analysis of the cellulose synthase-like genes *CSLD1*, *CSLD2*, and *CSLD4* in tip-growing *Arabidopsis* cells. Plant Physiol. 2008;148(3):1238–53.

66. Qin Y, Leydon AR, Manziello A, Pandey R, Mount D, Denic S, Vasic B, Johnson MA, Palanivelu R. Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. PLoS Genet. 2009;5(8):e1000621.

67. Xie B, Wang X, Hong Z. Precocious pollen germination in *Arabidopsis* plants with altered callose deposition during microsporogenesis. Planta. 2010;231(4):809–23.

68. Zhao L-N, Shen L-K, Zhang W-Z, Zhang W, Wang Y, Wu W-H. Ca²⁺–dependent protein kinase11 and 24 modulate the activity of the inward rectifying K+ channels in *Arabidopsis* pollen tubes. Plant Cell. 2013;25(2):649–61.

69. Morant M, Jørgensen K, Schaller H, Pinot F, Møller BL, Werck-Reichhart D, Bak S. CYP703 is an ancient cytochrome P450 in land plants catalyzing in-chain hydroxylation of lauric acid to provide building blocks for sporopollenin synthesis in pollen. Plant Cell. 2007;19(5):1473–87.

70. Dobritsa AA, Geanconteri A, Shrestha J, Carlson A, Kooyers N, Coerper D, Urbanczyk-Wochniak E, Bench BJ, Sumner LW, Swanson R, Preuss D. A large-scale genetic screen in *Arabidopsis* to identify genes involved in pollen exine production. Plant Physiol. 2011;157(2):947–70.

71. Xiong SX, Lu JY, Lou Y, Teng XD, Gu JN, Zhang C, Shi QS, Yang ZN, Zhu J. The transcription factors MS188 and AMS form a complex to activate the expression of *CYP703A2* for sporopollenin biosynthesis in *Arabidopsis thaliana*. Plant J. 2016;88(6):936–46.

72. Zhao D-Z, Wang G-F, Speal B, Ma H. The *EXCESS MICROSPOROCYTES1* gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the *Arabidopsis* anther. Genes Dev. 2002;16(15):2021–31.

73. Albrecht C, Russinova E, Hecht V, Baaijens E, de Vries S. The *Arabidopsis thaliana* SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 control male sporogenesis. Plant Cell. 2005;17(12):3337–49.

74. Mizuno S, Osakabe Y, Maruyama K, Ito T, Osakabe K, Sato T, Shinozaki K, Yamaguchi-Shinozaki K. Receptor-like protein kinase 2 (RPK 2) is a novel factor controlling anther development in *Arabidopsis thaliana*. Plant J. 2007; 50(5):751–66.

75. De Storme N, Geelen D. The *Arabidopsis* mutant *jason* produces unreduced first division restitution male gametes through a parallel/fused spindle mechanism in meiosis II. Plant Physiol. 2011;155(3):1403–15.

BMC
Plant Biology

RESEARCH ARTICLE

Open Access

Flower development, pollen fertility and sex expression analyses of three sexual phenotypes of *Coccinia grandis*

Amita G Ghadge[1], Kanika Karmakar[2], Ravi S Devani[1], Jayeeta Banerjee[1], Boominathan Mohanasundaram[1], Rabindra K Sinha[2], Sangram Sinha[2] and Anjan K Banerjee[1*]

Abstract

Background: *Coccinia grandis* is a dioecious species of Cucurbitaceae having heteromorphic sex chromosomes. The chromosome constitution of male and female plants is 22 + XY and 22 + XX respectively. Y chromosome of male sex is conspicuously large and plays a decisive role in determining maleness. Sex modification has been studied in hypogynous *Silene latifolia* (Caryophyllaceae) but there is no such report in epignous *Coccinia grandis*. Moreover, the role of organ identity genes during sex expression in *Coccinia* has not been evaluated earlier. Investigations on sexual phenotypes of *C. grandis* including a rare gynomonoecious (GyM) form and $AgNO_3$ mediated sex modification have added a new dimension to the understanding of sex expression in dioecious flowering plants.

Results: Morphometric analysis showed the presence of staminodes in pistillate flowers and histological study revealed the absence of carpel initials in male flowers. Though GyM plant had XX sex chromosomes, the development of stamens occurred in hermaphrodite flowers but the pollens were not fertile. Silver nitrate ($AgNO_3$) application enhanced stamen growth in wild type female flowers like that of GyM plant but here also the pollens were sterile. Differential expression of *CgPI* could be involved in the development of different floral phenotypes.

Conclusions: The three principle factors, Gynoecium Suppression (Su^F), Stamen Promoting Factor (SPF) and Male Fertility (m^F) that control sex expression in dioecious *C. grandis* assumed to be located on Y chromosome, play a decisive role in determining maleness. However, the characteristic development of stamens in hermaphrodite flowers of GyM plant having XX sex chromosomes indicates that Y-linked SPF regulatory pathway is somehow bypassed. Our experimental findings together with all other previous chromosomal and molecular cytogenetical data strongly support the view that *C. grandis* could be used as a potential model system to study sex expression in dioecious flowering plant.

Keywords: *Coccinia grandis*, Hypogynous, Epigynous, Dioecious, Gynomonoecy, Heteromorphic sex chromosomes, Sex modification, Organ identity genes, Silver nitrate

* Correspondence: akb@iiserpune.ac.in
[1]Indian Institute of Science Education and Research (IISER Pune), 900 NCL Innovation Park, Dr. Homi Bhabha road, Pune 411 008, Maharashtra, India
Full list of author information is available at the end of the article

Background

The vast majority of angiosperms are hermaphrodites having bisexual flowers and nearly 10% of the flowering plants produce unisexual flowers [1]. Sexual systems are coupled with the numerous combinations of unisexual and hermaphrodite flowers. There are about 6% angiosperms which are dioecious bearing male and female flowers on separate individuals [2,3]. Literature study suggests that dioecious plants have evolved independently and multiple times from their bisexual progenitors [4-6].

In comparison to animals, dioecious plants show relatively recent origin of sex chromosome evolution [7,8]. Sex determination in dioecious plants may be either genetically or environmentally controlled phenomenon [9]. Some dioecious plant species have fertile bisexual relatives [10], which are excellent system for sex chromosome study. The occurrence of sex chromosomes in dioecious plants is surprisingly rare and only 19 species are known to have heteromorphic sex chromosomes [10]. The heteromorphic sex chromosomes are well-studied in *Silene latifolia* (Caryophyllaceae), in which male and female plants carry XY and XX sex chromosomes respectively [11]. The Y chromosome is reported to be the largest of all chromosomes [12] and it consists of three sex determining regions viz., Gynoecium Suppression Factor (Su^F), Stamen Promoting Factor (SPF) and Male Fertility Factor (m^F) [13,14]. Other well-studied dioecious plants are *Rumex acetosa* exhibiting X to autosome ratio [15,16] and Poplar known for ZW system [17] for sex determination. In papaya, sex determination is controlled by a pair of recently evolved sex chromosomes, Y controlling male and Y^H controlling hermaphrodite [18]. Thus, sex chromosome study in different dioecious plant species provides an insight for better understanding of plant sex chromosome evolution.

Plant sex determination genes were so far identified from monoecious species by map based cloning approach because there is no recombination suppression at the sex determination loci [19]. Recent genomic technologies augmented the identification of X- and Y-linked genes and allowed the detection of dosage compensation of X- linked genes in *S. latifolia* [20-22]. In papaya, 8.1 Mb hermaphrodite-specific region of the Y^H chromosome (HSY) and its 3.5 Mb X chromosome counterpart were sequenced and annotated for identification of sex determination genes [23-25].

It is now well documented that silver nitrate (AgNO$_3$) as well as silver thiosulfate (Ag$_2$S$_2$O$_3$) have masculinizing effect on many dioecious and monoecious plants [26-29]. Beyer [30] reported that AgNO$_3$ acts as an anti-ethylene agent and induces male flowers by suppressing female reproductive organs. Evidences are also there that AgNO$_3$ can modify sex via inhibition of ethylene [29,31,32]. However, a study in *Silene latifolia*, contradicts this hypothesis and proposes that sex modification might be mediated by

inhibition of sulfahydryl enzymes upon application of silver thiosulfate [28]. Janousek *et al.* [33] showed that 5-azacytidine treated male plants of *S. latifolia* developed hermaphrodite flowers due to hypomethylation. This indicated the possible role of epigenetic control in sex determination and modification. Another unique case of sex modification is observed due to smut fungus (*Microbotryum violaceum)* infection in *Silene latifolia*. This fungus was reported to induce the development of anthers in female flowers (XX genotype) of *Silene latifolia* [34]. However, in this case, pollens were found sterile indicating the decisive role of Y chromosome in fertility of pollens. Investigations on sex modification in dioecious plants may enhance our knowledge on how a genetically controlled program gets modified to an altered state.

Unlike *Silene latifolia* (Caryophyllaceae), *Rumex acetosa* (Polygonaceae), *Carica papaya* (Caricaceae), *Spinacia oleracea* (Chenopodiaceae) and *Populus* (Salicaceae) [16,17,35,36], which have been well characterized to understand the mechanism of sex determination, *Coccinia grandis*, a member of Cucurbitaceae family having an inferior ovary received comparatively less attention. *Coccinia* is a small genus comprising 27 species, all dioecious in nature [37]. It is one of the few dioecious plant species, in which presence of heteromorphic sex chromosomes is reported. The chromosome constitution of male and female plants is 22 + XY and 22 + XX respectively [38]. Literature survey suggests that sexual dimorphism in *C. grandis* is determined by a large Y chromosome [38-41], which appears to be of comparatively recent origin [37]. However, the genes involved in sex determination of *C. grandis* are not yet known. Genome of *C. grandis* is almost six times smaller than that of *Silene latifolia* and is closely related to four fully sequenced genomes of Cucurbitaceae species [42,43]. Y chromosome of *C. grandis* is the largest one found in land plants; and it is heterochromatic, differently from the euchromatic Y chromosome of *S. latifolia* [43].

In addition to male and female sex forms of *C. grandis*, Kumar and Viseveshwaraiah [38] reported a gynodioecious form in which male flowers of the hermaphrodite plants were sterile. Earlier, Holstein and Renner [37] recorded a sexual phenotype of *C. intermedia* having male and female flowers/fruits on the same node. In the present investigation, we have identified a rare gynomonoecious plant (herein after referred as GyM), bearing hermaphrodite (GyM-H) and pistillate (GyM-F) flowers on the same plant. The presence of this naturally occurring GyM plant provides a great opportunity to study the genetic basis of sex determination in *C. grandis*.

To understand the floral development and sex expression in *C. grandis*, we aimed at a comprehensive characterization of sexual phenotypes through morphometric, histological, chromosomal and molecular approaches. In the present investigation, it was observed that foliar spray

of AgNO$_3$ is able to induce hermaphrodite flowers in wild type female plants. To determine whether Organ Identity Genes (OIGs) have any role in differentiation of the sexes, expression studies were carried out in male, female and GyM plants. To our knowledge, no such report for *C. grandis* is available in the literature.

Results

Morphological differences amongst three sexual phenotypes

While there exist striking similarities in inflorescence, sepal and petal characters, differences in the morphology of mature flowers were clearly observed amongst the three sexual phenotypes. Mature male flowers were seen to be composed of three whorls having five sepals, five united petals and five (2 + 2 + 1) synandrous stamens (Figure 1A,E). In contrast, the female flowers were composed of four whorls. While sepals and petals were identical to male flowers, the stamens were found to be arrested as rudimentary staminodes. The gynoecium consisted of three carpels having a fused style with three bifid stigmas (Figure 1B,F). The GyM plants bear two different types of flowers (i)

Figure 1 Morphology of mature flowers of *Coccinia grandis*. Macroscopic view of staminate flower **(A)** of male plant, pistillate flower **(B)** of female plant, hermaphrodite (GyM-H) **(C)** and pistillate (GyM-F) **(D)** flowers of gynomonoecious (GyM) plant with petals cut open. Petals removed from staminate flower **(E)** of male plant, pistillate flower **(F)** of female plant, hermaphrodite (GyM-H) **(G)** and pistillate (GyM-F) **(H)** flowers of gynomonoecious (GyM) plant to show inner floral organs. st: Stamens, c: carpels, rst: rudimentary stamens, o: ovary. Scale bars =1 cm.

hermaphrodite (GyM-H) and (ii) pistillate (GyM-F) (Additional file 1: Figure S1). The GyM-H flowers had four whorls, almost similar to the flowers of female sex; the only difference being here that the staminodes gradually developed to mature stamens (Figure 1C,G). It was also observed that some of the GyM-H flowers exhibited incomplete growth of stamens (Additional file 2: Figure S2A) as well as petaloid stamens (Additional file 2: Figure S2C). The organization of floral organs in GyM-F flowers of GyM plant was found to be similar to that of wild type female plant (Figure 1D,H). We observed random positional distribution of GyM-F and GyM-H flowers in GyM plant and the ratio of these flowers was found to be approximately 30:70 during the months of April to July. The phylogenetic analysis using *mat*K and *trn*S^{GCU}-*trn*G^{UCC} intergenic spacer region, revealed that the GyM plant is another sexual phenotype of *Coccinia grandis* (Additional file 3: Figure S3). Except for the three sexual phenotypes of *Coccinia grandis* (Additional file 4: Table S1), sequences for constructing the phylogenetic tree were used from the previously published data [37]. Seed content of fruits from female plant (seed number and seed weight per fruit) was observed to be higher than that of fruits from GyM plant (Additional file 5: Figure S4A,B).

Histological analysis

To understand the sequential development of sex organs, histological analysis was carried out at different stages of flower development for all three sexual phenotypes (Figure 2A–T).

Male: Histological observation of male flowers (stages 3–4, Additional file 6: Figure S5A) showed the presence of sepals, petals and stamens having no sign of carpel initials (Figure 2A). Even in the later stages of flower development, any rudimentary carpel was not observed. However, the possibility of presence of carpel initials in primordial stages of flower development cannot be completely ruled out. Further growth of stamens was observed in the successive stages of male flower development (Figure 2B–D). Finally, in stage 12 (Additional file 6: Figure S5A), mature pollens were found inside the anthers when petals were about to open (Figure 2E, Additional file 7: Figure S6).

Female: Whereas female flowers (stages 3–4, Additional file 6: Figure S5B) exhibited the presence of sepals, petals, stamen initials and carpels having an inferior ovary in four whorls (Figure 2F). While development of the androecium remained arrested in early stages, growth of the gynoecium was noted in successive stages of development (Figure 2G–I). At stage 12 (Additional file 6: Figure S5B), when the petals were about to open, the gynoecium was found to be completely developed (Figure 2J).

GyM: Presence of sepals, petals, stamens and carpel initials along with an inferior ovary was observed in four
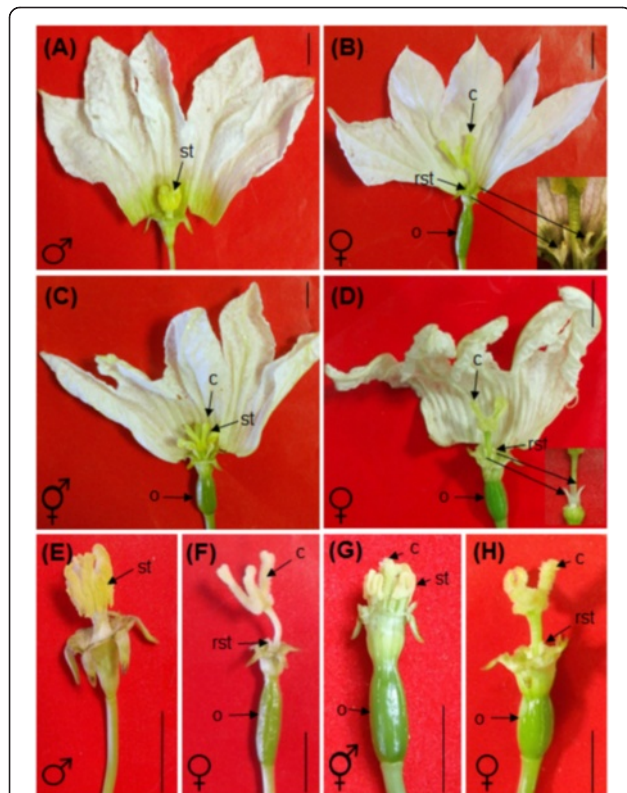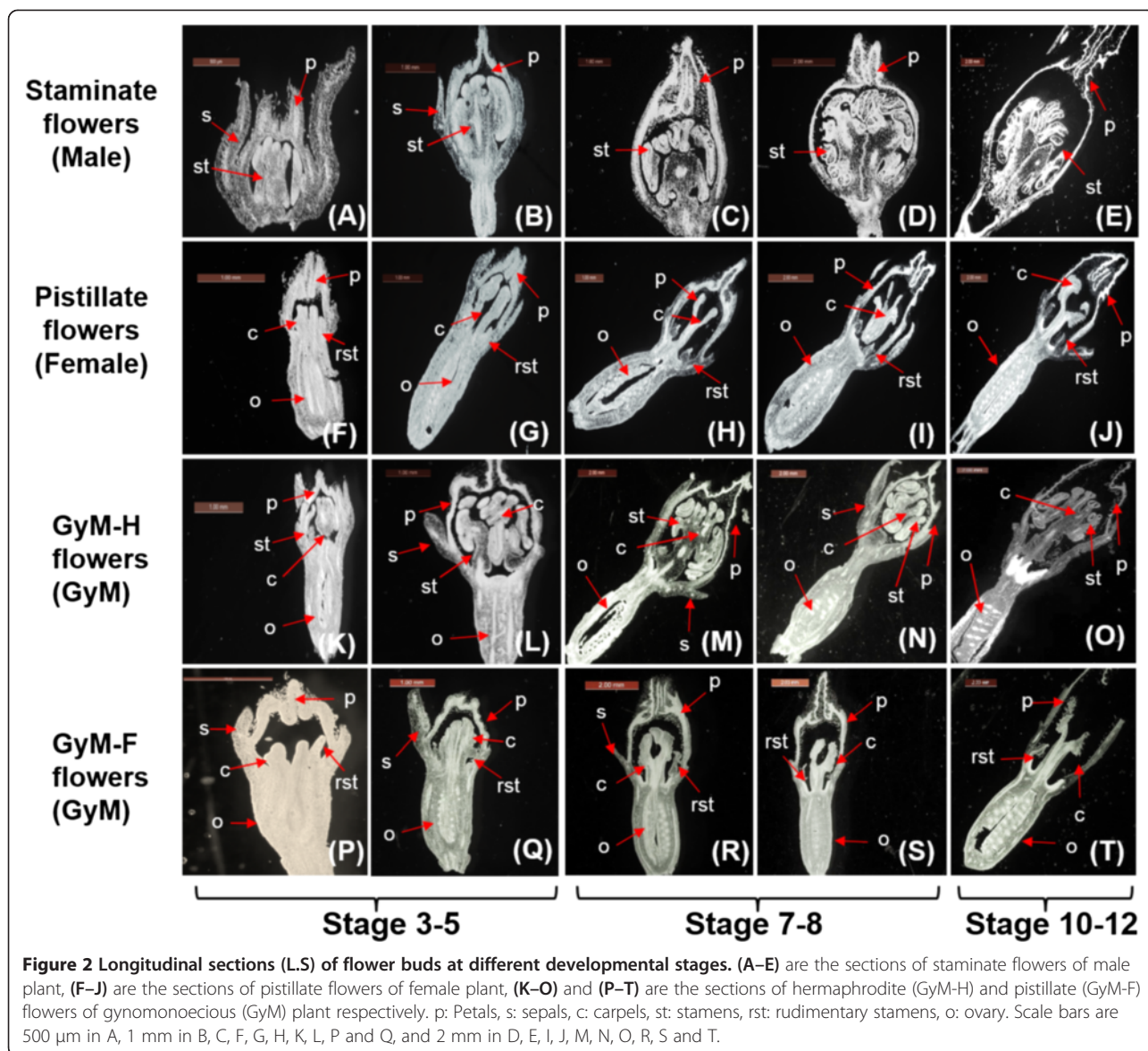
Figure 2 Longitudinal sections (L.S) of flower buds at different developmental stages. (A–E) are the sections of staminate flowers of male plant, **(F–J)** are the sections of pistillate flowers of female plant, **(K–O)** and **(P–T)** are the sections of hermaphrodite (GyM-H) and pistillate (GyM-F) flowers of gynomonoecious (GyM) plant respectively. p: Petals, s: sepals, c: carpels, st: stamens, rst: rudimentary stamens, o: ovary. Scale bars are 500 μm in A, 1 mm in B, C, F, G, H, K, L, P and Q, and 2 mm in D, E, I, J, M, N, O, R, S and T.

successive whorls of GyM-H flowers at early stages of development (Figure 2K, Additional file 2: Figure S2B). Further growth of the gynoecium and androecium occurred in successive stages of development (Figure 2L–N) and at stage 12 (Additional file 6: Figure S5C), growth of the gynoecium and androecium was found to be complete (Figure 2O). However, development of GyM-F flowers in GyM plant was found to be identical to that of wild type female plant (Figure 2P–T).

Chromosomal study

In order to have a better understanding of the relation between male, female and GyM plants of *C. grandis* growing in the same environment, comparative cytological studies were carried out. The somatic chromosome number of male, female and GyM plant was found to be $2n = 24$

(Table 1). Sex chromosomes were heteromorphic and in male plants Y chromosome was conspicuously large (Figure 3A). In wild type female and GyM plants, the chromosome constitution is $22 + XX$ (Figure 3B,C). The karyotype of wild type female and GyM plant showed similarity to a considerable extent (Figure 3B,C). Meiotic studies of male sex showed end to end pairing between X and Y chromosomes (Figure 3D). In contrast, normal pairing of homologous chromosomes were found in GyM-H flowers of GyM plant (Figure 3E).

AgNO₃ induced sex modification

Different concentrations of silver nitrate (AgNO₃) solution were sprayed on the basal leaves of male, female and GyM plant (Additional file 8: Table S2). Newly emerging flower buds of wild type female plants showed

Table 1 Numerical data on somatic chromosome complements of *C. grandis* (male, female and gynomonoecious (GyM) plants)

Chromosome numbers	Chromosome size (µm)* (Mean ± SD)			F%			Position of centromere		
	Male	Female	GyM	Male	Female	GyM	Male	Female	GyM
1	1.92 ± 0.07	2.01 ± 0.03	1.92 ± 0.06	50	50	50	m	m	m
2	1.92 ± 0.06	1.92 ± 0.06	1.92 ± 0.07	45	45	48	nm	nm	nm
3	1.76 ± 0.03	1.84 ± 0.06	1.82 ± 0.04	50	50	50	m	m	m
4	1.76 ± 0.03	1.76 ± 0.03	1.76 ± 0.06	45	47	45	nm	nm	nm
5	1.62 ± 0.07	1.62 ± 0.03	1.65 ± 0.03	33	33	33	sm	sm	sm
6	1.62 ± 0.03	1.62 ± 0.03	1.65 ± 0.03	46	46	46	nm	nm	nm
7	1.54 ± 0.09	1.54 ± 0.07	1.54 ± 0.07	43	43	43	nm	nm	nm
8	1.54 ± 0.03	1.54 ± 0.07	1.54 ± 0.09	46	47	47	nm	nm	nm
9	1.40 ± 0.02	1.54 ± 0.07	1.43 ± 0.05	44	47	47	nm	nm	nm
10	1.22 ± 0.09	1.22 ± 0.05	1.23 ± 0.03	44	46	46	nm	nm	nm
11	1.22 ± 0.01	1.22 ± 0.05	1.23 ± 0.03	45	45	46	nm	nm	nm
12	1.10 ± 0.06	1.10 ± 0.06	1.10 ± 0.06	44	46	47	nm	nm	nm
Y[1]	4.60 ± 0.07	-	-	48	-	-	nm	-	-

*Mean of 5 metaphase plates. GyM: gynomonoecious, m: metacentric, nm: nearly metacentric, sm: submetacentric. The karyotype of male and female plants was compared with the gynomonoecious (GyM) chromosomes. Y[1]: Single Y chromosome present in male sex.

enhanced growth of stamens after application of $AgNO_3$ solution (Figure 4A–D) whereas; male flowers did not show any changes in floral structure. Histological studies further confirmed the dose dependent stamen growth in wild type female flowers (Figure 4H–K; Additional file 8: Table S2). However, concentrations higher than 35 mM

had lethal effect. At dosages of 30 and 35 mM of $AgNO_3$, the morphology of newly developed flowers was comparable to GyM-H flowers after 10-12 days of observation (Figure 4D–G). Interestingly, all mature flowers in GyM plant were found to be hermaphroditic after application of $AgNO_3$, indicating that even the staminodes

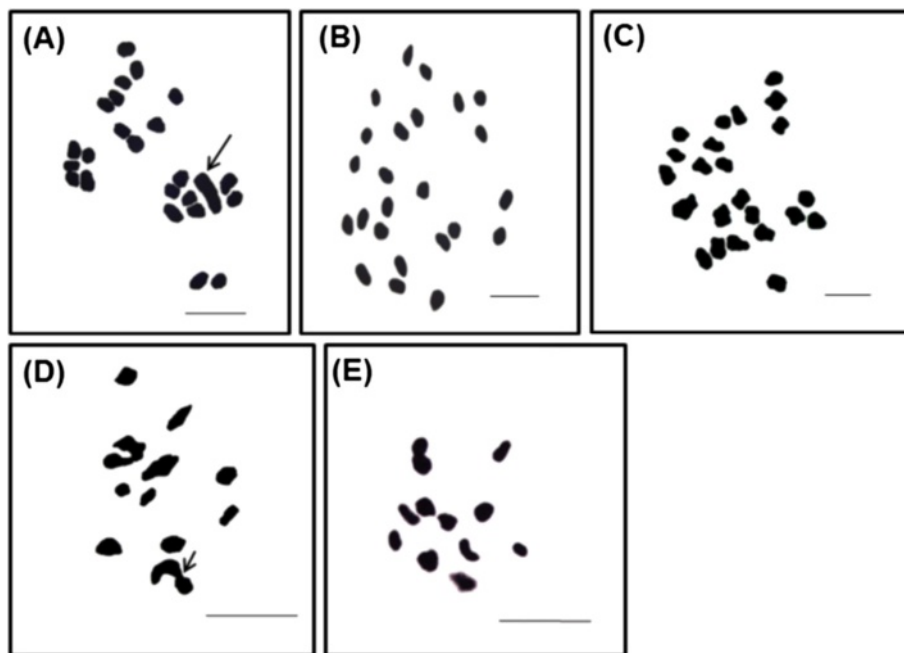**Figure 3 Metaphase chromosomes of *C. grandis*.** Mitotic metaphase chromosomes showing 2*n* =24 chromosomes of male **(A)** (arrow indicates the large Y chromosome), female **(B)** and gynomonoecious (GyM) **(C)** plants. Meiotic metaphase chromosomes showing 12 bivalents of male **(D)** (arrow indicates end to end pairing of X and Y chromosomes), gynomonoecious (GyM) **(E)** plants. Scale bar =5 µm.
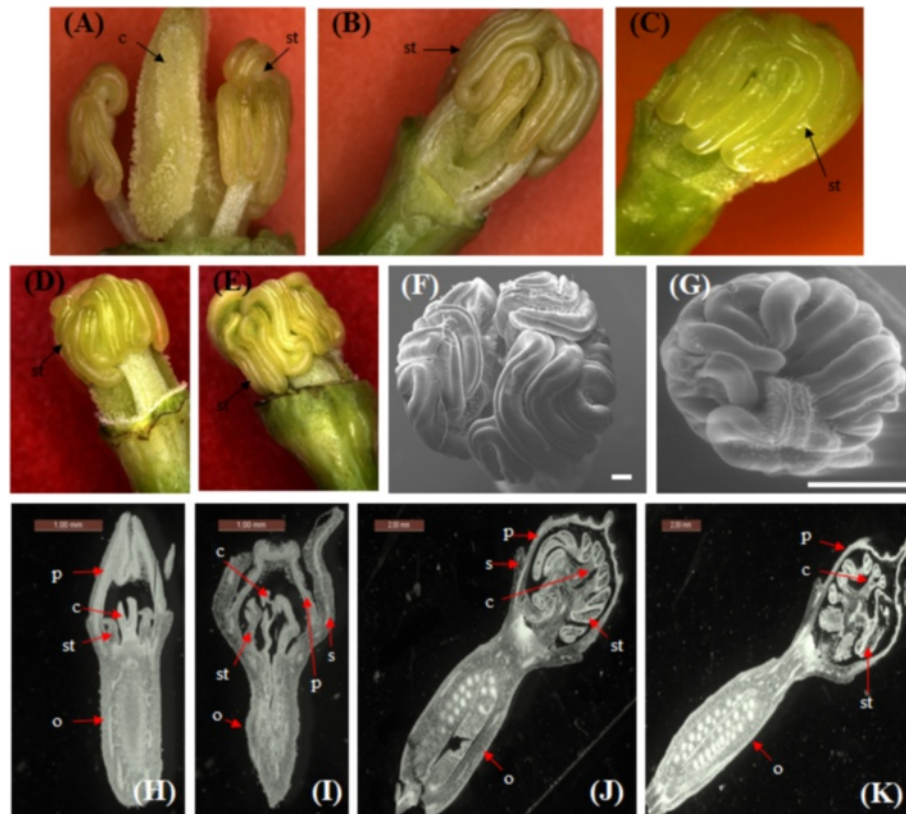
Figure 4 Effects of silver nitrate (AgNO$_3$) solution on female plant. (A-C) are the pictures of female flowers after spraying of AgNO$_3$ solution showing gradual enhanced stamen growth. Magnified view of stamens in **(D)** pistillate flowers of AgNO$_3$ treated female plant and **(E)** hermaphrodite (GyM-H) flowers of gynomonoecious (GyM) plants. Scanning electron micrographs of top view of **(F)** pistillate flowers from AgNO$_3$ treated female plant and **(G)** hermaphrodite (GyM–H) flowers of gynomonoecious (GyM) plants. Petals and sepals have been removed to better view sexual structures. Longitudinal sections **(H-K)** of flower buds of silver nitrate treated female plant (after spraying of 35 mM silver nitrate solution). H, I – flower buds of stage 5, J – flower bud of stage 8 and K – flower bud of stage 10. p: Petals, s: sepals, c: carpels, st: stamens, o: ovary. Scale bars are 300 μm in F, 1 mm in G, H and I, and 2 mm in J and K.

of pistillate flower buds have developed into mature stamens (Additional file 9: Figure S7).

Mating experiments and pollen fertility

Mating experiments were designed to investigate the fertility of pollens from male flowers and GyM-H flowers (Table 2). The crosses between male and emasculated GyM-H resulted in 83.33% of fruit setting. No fruit setting was recorded in crosses between GyM-H and wild type female flowers. It was also noted that 90% fruit setting occurred in crosses between male and wild type female (Table 2). Similarly, the crosses between the wild type male and the pistillate flowers of GyM plant also yielded 93% of fruit setting. However, no fruit setting was achieved in crosses between GyM-H and GyM-F flowers and by selfing GyM-H (Table 2).

For viability assays, pollens were isolated from opened flowers of male, GyM-H and converted flowers of AgNO$_3$ treated female plant. Pollens from male flowers took aceto-carmine stain; whereas pollens from GyM-H flowers and

converted flowers of AgNO$_3$ treated female plant did not retain any stain (Figure 5A–C). These results were reconfirmed with FDA test (Figure 5D–F). In addition, pollen germination was also tested for male, GyM plant and AgNO$_3$ treated female plant. Highest frequency of pollen germination (38%) was achieved when pollens of male flowers were incubated in 5% sucrose solution containing required amount of Ca(NO$_3$)$_2$ and H$_3$BO$_3$ (Figure 5G,H). In contrast, pollens of hermaphrodite flowers of GyM and AgNO$_3$ treated female plant did not show any germination when incubated in different germinating media. From the above results, we concluded that pollens of male flowers are fertile and pollens from GyM-H and converted flowers of AgNO$_3$ treated female plant are sterile in nature.

Identification and expression analysis of Organ Identity Genes (OIGs)

In order to understand whether B and C class Organ Identity Genes (OIGs) have any role in determining the sex of the developing flowers of male, female and GyM plant,

Table 2 Mating design and percentage of fruit set in *C. grandis*

Mating design	Pollen source	No. of fruit set	% fruit set	Remarks
Male X GMH (emasculated)	Male	8.33 ± 0.577	83.33	Fertile pollen
Male X GMF	Male	9.33 ± 0.577	93.33	Fertile pollen
Male X Female	Male	9.0 ± 1.00	90.00	Fertile pollen
GMH self	GMH	0.00	0.00	Sterile pollen
GMH X GMF	GMH	0.00	0.00	Sterile pollen

Replications =3, N =30, No. of crosses/ mating design are 10 for all the above sets.
GyM-H: hermaphrodite flower from gynomonoecious (GyM) plant, GyM-F: pistillate flower from gynomonoecious (GyM) plant.

CgPI (a B class OIG) and *CgAG* (a C class OIG) were isolated and an expression analysis was carried out using quantitative real-time PCR (qRT-PCR). The degenerate primers based on the conserved amino acid sequences of PI (*PISTILLATA*) and AG (*AGAMOUS*), yielded ~350 bp of *PISTILLATA* (*CgPI*) and ~250 bp of *AGAMOUS* (*CgAG*) homologs through RT-PCR reaction. The partial sequences

for *CgPI* [DDBJ:AB859715] and *CgAG* [DDBJ:AB859714] have been deposited in DDBJ. Full length transcript sequences were deduced from 5′ and 3′ RACE products and amplicons of *CgPI* (~893 bp) and *CgAG* (~952 bp) were obtained (Figure 6A). cDNA for *CgPI* and *CgAG* coded for putative proteins of 212 and 232 amino acids respectively. The deduced amino acids sequences for both the genes
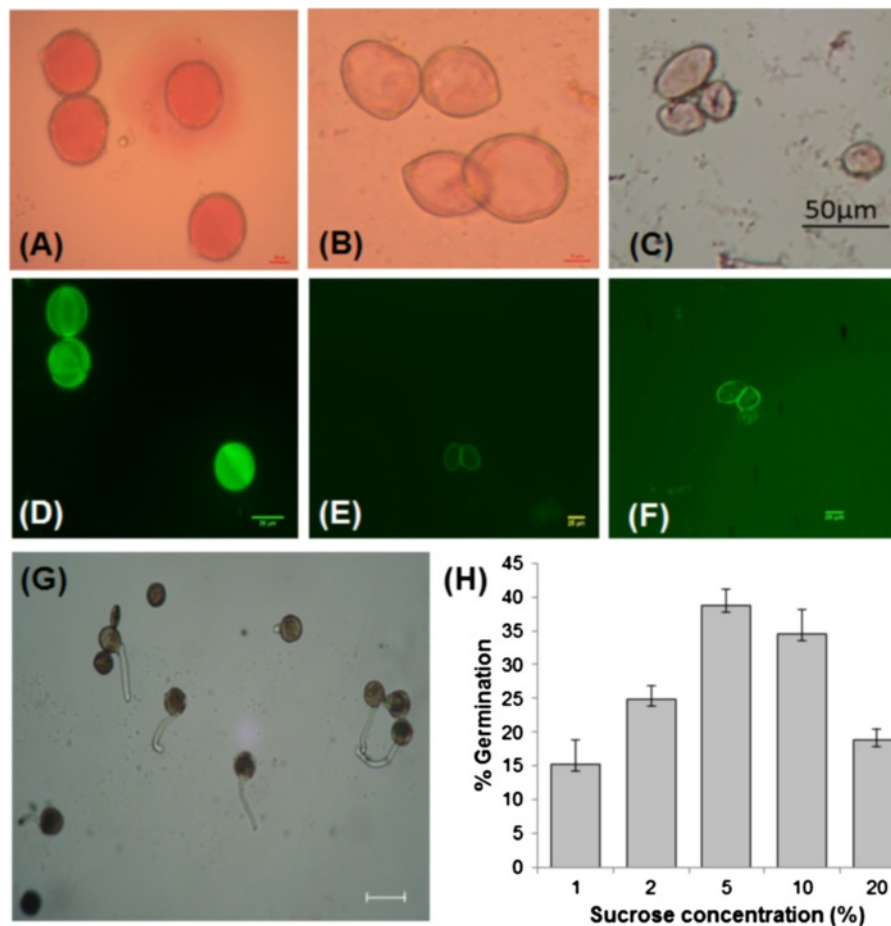
Figure 5 Viability tests of pollens from male, gynomonoecious (GyM) and AgNO₃ treated female plants. Pollens stained with 1% acetocarmine from male **(A)**, gynomonoecious (GyM) **(B)** and AgNO₃ treated female **(C)** plants. **(D)**, **(E)** and **(F)** are the fluorescein diacetate (FDA) stained pollens from male, gynomonoecious (GyM) and AgNO₃ treated female plants respectively. Pollens stained with acetocarmine **(A)** and FDA **(D)** are viable. Scale bars are 10 µm in A, 5 µm in B, 50 µm C, and 25 µm in D, E and F. **(G)** Highest germination of male pollens in 5% sucrose solution. Scale bar =50 µm. **(H)** Graphical representation of the germination percentage in different concentrations of sucrose solutions. Means ± standard errors are reported in the graph; n = 10.

Figure 6 (See legend on next page.)

(See figure on previous page.)

Figure 6 Full length *CgPI* and *CgAG* transcript isolation and multiple sequence alignment of deduced amino acid sequences.
(A) Amplification of full length *CgPI* and *CgAG* transcripts from total RNA harvested from flower buds. **(B)** Comparison of CgPI with other PISTILLATA-like genes. **(C)** Comparison of CgAG with other AGAMOUS-like genes. Conserved region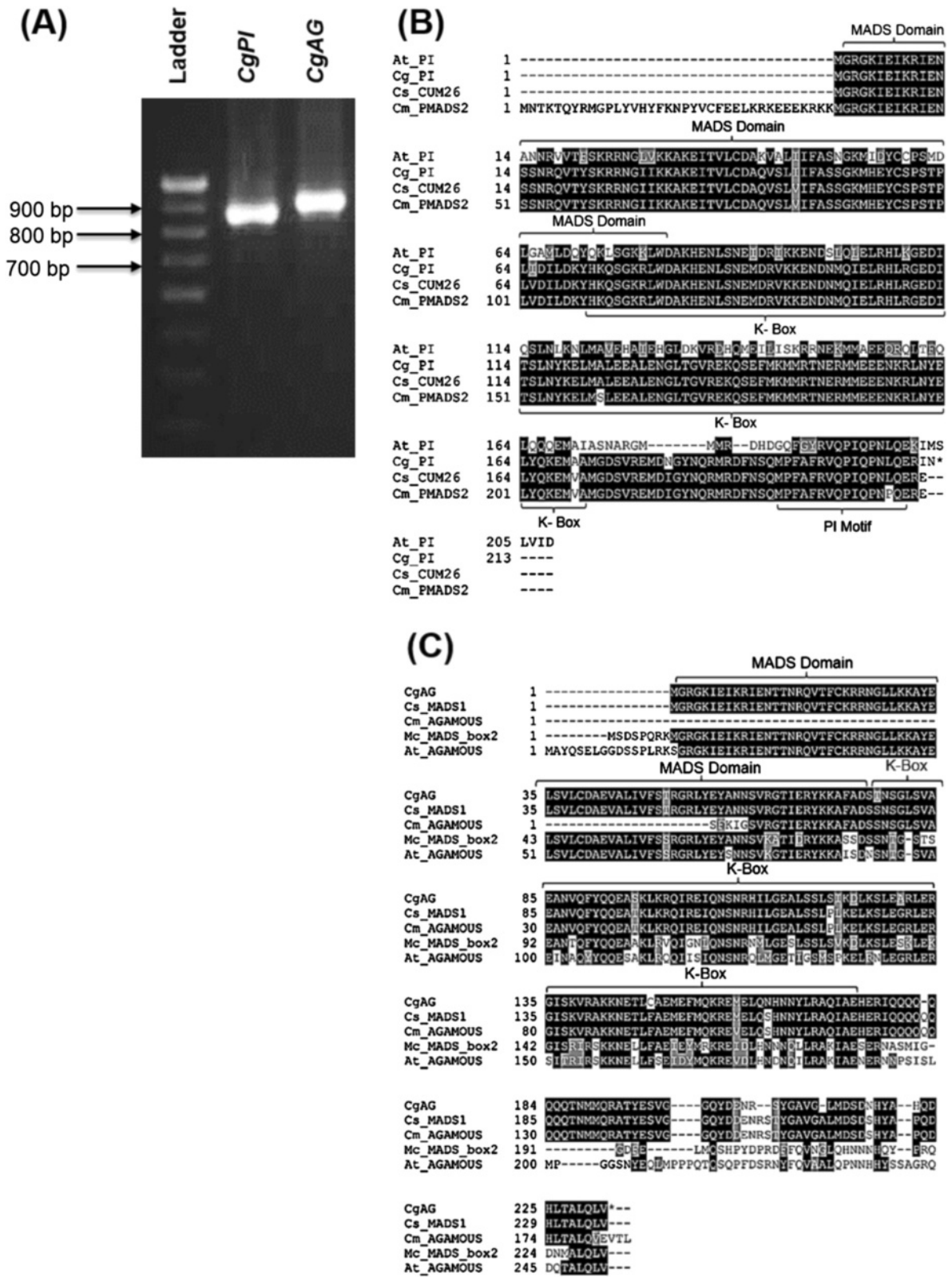s are shaded in black. At_PI, Cg_PI, Cs_CUM26 and Cm_pMADS2 are *PISTILLATA* like genes from *Arabidopsis thaliana*, *Coccinia grandis*, *Cucumis sativus* and *Cucumis melo* respectively. Cg_AG, Cs_MADS1, Cm_AGAMOUS, Mc_MADS_box2, At_AGAMOUS are *AGAMOUS* like genes from *Coccinia grandis*, *Cucumis sativus*, *Cucumis melo*, *Momordica charantia* and *Arabidopsis thaliana* respectively. MADS domain and K-box are identified by NCBI's conserved domain database and marked accordingly.

showed high conservation when aligned with other *PISTILLATA* and *AGAMOUS* like genes (Figure 6B,C). Two consensus regions, MADS domain and K-box were found on the deduced amino acid sequences (Figure 6B,C).

CgPI, a B class gene required for petal and stamen development, was found to be expressed in male, wild type female and GyM flower buds (Figure 7A). Expression of *CgAG*, a C class gene essential for stamen and carpel development, was also noted in male, wild type female and GyM flower buds (Figure 7B). Our results showed that both these genes are expressed in all developmental stages (Additional file 6: Figure S5) (early, middle and late) of flowers from male, female and GyM plant. *CgPI* had a significant difference of expression across all three sexual forms during early, middle and late developmental stages (Figure 7A), while *CgAG* showed significant differential expression in buds of early stages only (Figure 7B). We have also noted that *CgPI* expression is comparatively high in male flower buds than that of wild type female buds. However, GyM flowers exhibited an intermediate level of *CgPI* expression in early and late staged buds (Figure 7A). Further, our results for stamen-specific expression analysis showed a significant difference for both *CgPI* and *CgAG* levels between stamens of male, GyM-H, AgNO$_3$ treated female plant, rudimentary stamens of GyM-F and wild type female plant (Figure 7C,D). Surprisingly, rudimentary stamens of GyM-F showed higher *CgPI* expression than stamens of GyM-H flowers (Figure 7C).

Discussion
Carpel and stamen differentiation programmes follow independent pathway
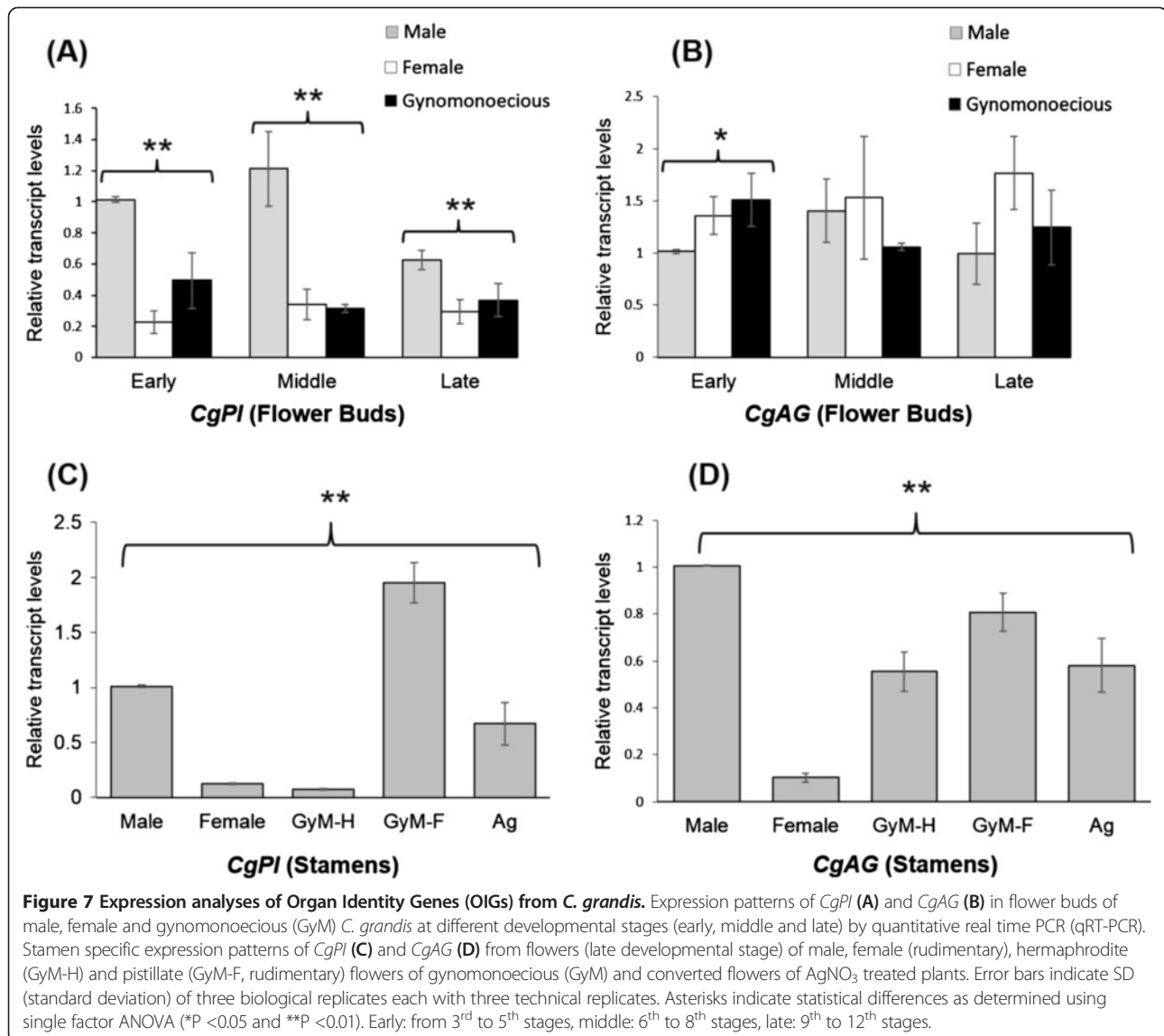In contrast to *Silene latifolia*, where rudimentary gynoecium is found in male flower [44,45] histological study revealed the absence of carpel initials even at early stage of development (stages 3-4, Additional file 6: Figure S5A) of male flower in *C. grandis* (Figure 2A,B). Though stamen initiation occurs in female plants, its growth is arrested at early stages (stages 4–5, Additional file 6: Figure S5B) of flower development (Figure 2G–I) leading to the retention of sterile staminode in mature flower. This indicates a functional interference in the male differentiation pathway of female flowers as was reported in *Silene latifolia* [14]. In GyM-H flowers, androecium and gynoecium develop simultaneously till maturation (Figure 2K, L) and arrest of stamen or carpel growth is not observed (Figure 2M–O).

However, in pistillate flowers of GyM plant, arrest of stamen growth occurs at early stages like the flowers of wild type female plant (Figure 2Q,R). The development of mature carpel with arrested stamen growth as evidenced by the presence of rudimentary staminodes in pistillate (GyM-F) flowers and the synandrous stamens with fully grown carpel in GyM-H flowers indicate that the carpel and stamen differentiation programmes follow independent pathway.

Gynomonoecious (GyM) *C. grandis* - is not a Y-deletion mutant
While investigating the morphological differences between male and female sexes, we have recorded the existence of a GyM plant in the north eastern part of India (Tripura) that exhibited morphological characteristics similar to that of male and female sex forms of *C. grandis*. The morphological characterization and the phylogenetic analysis, based on the tree constructed with *mat*K and *trn*S^{GCU}-*trn*G^{UCC} intergenic spacer regions clearly establish the identity of the GyM plant to be another sexual phenotype of *C. grandis*.

The present record of diploid chromosome number $2n = 24$ in both male and female sexes (Figure 3A,B) and the presence of heteromorphic sex chromosomes in male plants corroborate previous findings and validate XY sex determination system [38,43,46-48]. The characteristic end to end pairing between X and Y chromosomes (Figure 3D) indicates recombination between Pseudo Autosomal Region (PAR) [43] and that there are non-recombining regions between X and Y chromosomes as was suggested by other researchers to explain the genetic basis of sex determination in some dioecious plants [13,49]. The absence of carpel initial in male plant suggests that the Y chromosome has a dominant gynoecium suppressor gene at the non-recombining region like that of *S. latifolia* [13]. The karyotype of GyM plant shows high degree of similarity to that of wild type female (Table 1; Figure 3B,C). The smallest bivalent found in metaphase I of hermaphrodite flower does not match with the size of X chromosome of heteromorphic pair found in male sex (Figure 3D,E). Therefore, it requires further test to assume the smallest chromosomes as X chromosome [43] and at this stage, it remains inconclusive due to the unavailability of X- specific probes in *C. grandis*. The absence of male specific Y chromosome in GyM plant and normal pairing between homologous chromosomes (Figure 3C,E) indicate that GyM plant also

Figure 7 Expression analyses of Organ Identity Genes (OIGs) from C. grandis. Expression patterns of *CgPI* **(A)** and *CgAG* **(B)** in flower buds of male, female and gynomonoecious (GyM) *C. grandis* at different developmental stages (early, middle and late) by quantitative real time PCR (qRT-PCR). Stamen specific expression patterns of *CgPI* **(C)** and *CgAG* **(D)** from flowers (late developmental stage) of male, female (rudimentary), hermaphrodite (GyM-H) and pistillate (GyM-F, rudimentary) flowers of gynomonoecious (GyM) and converted flowers of AgNO$_3$ treated plants. Error bars indicate SD (standard deviation) of three biological replicates each with three technical replicates. Asterisks indicate statistical differences as determined using single factor ANOVA (*P <0.05 and **P <0.01). Early: from 3rd to 5th stages, middle: 6th to 8th stages, late: 9th to 12th stages.

possesses 22 + XX chromosomes and contains genetic information necessary to produce both pistillate and hermaphrodite flowers. The lability of the expression of hermaphroditism suggests that the GyM plant is genotypically a female individual and not a Y- deletion mutant. The questions that might arise are firstly, in absence of the Y chromosome, how does the development of stamens occur in hermaphrodite (GyM-H) flowers of GyM plant? And secondly, what factors contribute to the development of hermaphrodite (GyM-H) and pistillate (GyM-F) flowers in the same plant?

Factors stimulating stamen development in GyM plant in absence of Y chromosome

In contrast to fertile and viable pollens of male flowers (Figure 5A,D), pollens of GyM-H flowers are sterile in nature and remain immature even at later stages of development (Figure 5B,E). The results of breeding experiments negate the possibility of self-fertilization and thus fruit setting occurs only through allogamy or cross pollination when pollens from male sex act as donor (Table 2). This indicates that viable and fertile pollens are produced in male plants only and that male fertility factor is located on Y chromosome. Evidently, male fertility is controlled by the Y chromosome and it plays a decisive role in determining sex in *C. grandis* [50]. Similar to *Silene latifolia* [13,14], our experimental results suggest that in *C. grandis*, at least three key factors: Gynoecium Suppressor, Stamen Promoting Factor And Male Fertility Factor have assembled and possibly rearranged during evolution of the Y chromosome. However, development of the stamen with sterile pollens in

GyM-H flowers of genotypically female GyM plant suggests that the factors stimulating stamen development might be present elsewhere in the genome, while the male fertility factor may be absent. Scutt et al. [51] reported that infected female S. latifolia with XX sex chromosomes can develop morphologically normal stamens. Whereas, Farbos et al. [14] have shown that in Silene latifolia, Gynoecium Suppression Factor (GSF/Su^F) and Stamen Promoting Factor (SPF) regions of Y-chromosome behave as linked dominant traits and that SPF is absent in female plants with XX sex chromosomes. In the absence of Y or truncated Y chromosome, the mechanism of stamen development in GyM-H flower of C. grandis is not clear but needs further investigation.

Phenomenon of silver nitrate induced stamen development resembles that of GyM plant

Silver nitrate stimulated stamen development in female plants of C. grandis mimics the pathway of stamen development in GyM plant (Figure 4D–G). Y chromosome is absent in both of these sexual phenotypes and pollens of converted flowers of AgNO$_3$ treated female plant are sterile in nature like the pollens of GyM-H flowers (Figure 5C,F). This suggests that stamen development is induced in wild type female by an unknown pathway which is independent of Y-mediated mechanism as was reported in Silene latifolia [28]. But there is no clue how this signal is transmitted from leaves to flowers that leads to sex modification. However, silver nitrate effect is transient and normal female flower develops after a period of 15-20 days. This may be due to the fact that the effective AgNO$_3$ concentration below the threshold level cannot impede the molecular mechanism leading to the formation of gynoecium with arrested stamen growth. It appears that AgNO$_3$ at an optimum concentration stimulates stamen development in wild type female and GyM-F of C. grandis possibly by a temporary delay in functional interference of male differentiation pathway. In such a condition, possibility of the presence of male repressive factor in untreated plants and its de-repression by AgNO$_3$ molecule in treated wild type female and GyM-F cannot be ruled out.

Differential expression of OIGs and the development of the three different floral phenotypes

The B and C function genes viz. CgPI and CgAG show homology to CUM26 and MADS1 of Cucumis sativus respectively. qRT-PCR studies suggest that like Silene latifolia [52], the male flowers of C. grandis had higher CgPI expression compared to wild type female flowers (Figure 7A). This observation was also true for stamen specific expression analysis (Figure 7C). The high expression of CgPI in stamens of GyM-F and reduced expression in stamens of GyM-H flowers cannot be explained currently and would require future investigations. This study indicates that OIGs might be under differential regulation in male, female

and GyM plant leading to the development of male, wild-type female and GyM-H as well as GyM-F flowers. To this effect, further studies are required to understand the role of ACS (aminocyclopropane-1-carboxylate synthase) and WIP1 (Wound Inducible Protein 1) genes, which were shown to govern sex expression in a related species, Cucumis melo [53,54]. The sex-determining locus A of melon encodes an ethylene biosynthesis enzyme, CmACS-7, that represses stamen development in female flowers. The G locus of melon encodes CmWIP1, a transcription factor that represses carpel development in male flowers. Also, it has been shown that the role of ACS gene is conserved in another member of the family Cucumis sativus [55]. Future investigation on functional validation of these genes would be necessary to decipher their role in sex expression and modification.

Conclusions

There is no doubt that the 'female suppressing' functions of the Y chromosome in male C. grandis is an initial event towards the establishment of the sexual dimorphism. The process of stamen initiation occurs in wild type female and GyM plants even in the absence of Y chromosome but the arrest of further development of stamens suggests a possible interference in 'Stamen Promoting' functions (SPF). The pollens of GyM-H and converted flowers of AgNO$_3$ treated female plants were sterile indicating that the male fertility factor is located on Y chromosome which is solely responsible for pollen fertility. The significance of GyM plant of C. grandis lies in its ability to develop stamens with sterile pollens because such evidences were not reported in any other plants including gynomonoecious Silene species [56,57]. The characteristic development of stamens in hermaphrodite flowers of GyM having XX sex chromosomes and AgNO$_3$ modified wild type female flowers is mediated by an unknown mechanism bypassing the Y-linked SPF regulatory pathway. Our experimental findings together with all other previous chromosomal and molecular cytogenetical data strongly support the view that C. grandis could be used as a potential model system to study sex expression in dioecious flowering plant.

Methods

Plant material and stages of flower development

Tuberous roots of wild type male, female and GyM Coccinia grandis were collected from west Tripura and grown in the experimental fields of IISER Pune and Tripura University (Herbarium voucher for gynomonoecious C. grandis is provided in Additional file 10: Figure S8). The clones were maintained in the experimental plots since last two years. Leaves and flowers from male, female and GyM plants were harvested periodically and were frozen in liquid nitrogen for various experimental

purposes. Based on the size of the flower buds, we have divided the process of flower initiation into 12 different stages in ascending order. Out of the 12 different stages, first two stages were studied under stereomicroscope and only the flower buds from 3rd to 12th stages (Additional file 6: Figure S5) were considered for stage specific histological study. For qRT-PCR expression analyses, flower buds were grouped into three different categories viz., early (from 3rd to 5th stages), middle (6th to 8th stages) and late (9th to 12th stages) for experimental purposes. In addition, stamens were also harvested from the flowers (late developmental stage) of male, GyM-H, and converted flowers of AgNO$_3$ treated female plant as well as rudimentary stamens of wild type female and GyM-F flowers for expression analysis.

Histology of flower buds

To understand the patterns of flower development of male, female and GyM plants, flower buds of different stages (Additional file 6: Figure S5) were harvested and fixed in 1:3 acetic acid-ethanol solution and kept at 4°C for overnight. Longitudinal sections (L.S.) of flower buds of different developmental stages were prepared as described by Cai and Lashbrook [58] with the following modifications. The fixed tissue was dehydrated with 75% ethanol for 40 min, 95% ethanol for 40 min and finally washed thrice with 100% ethanol, each with 45 min intervals. The material was then treated with 50-50% ethanol-xylene for 45 min, followed by clearing with 100% xylene for 45 min. Xylene was replaced with paraplast wax at 59°C. Then the tissue was embedded in the paraplast blocks. Thin paraplast sections (10 μm) were mounted on the slides with water at 50°C. The wax was cleared from the slides by washing with 100% xylene. The images of the cleared slides were finally documented in Leica MZ 16 FA microscope.

Analysis of mitosis and meiosis chromosomes

To analyze mitotic and meiotic chromosomes, investigations were carried out through modified aceto-orcein and aceto-carmine staining techniques [59]. Young leaf-tips were pre-treated in saturated solution of para-dichlorobenzene for 5 h at 10–15°C followed by overnight fixation in 1:3 acetic acid-ethanol mixture. The fixed leaf tips were then hydrolyzed with 5 N HCl at 10°C for 15 min, stained with 2% aceto-orcein for overnight and finally squashed in 45% acetic acid. For meiotic chromosome preparation, young flower buds were fixed in 1:3 acetic acid-ethanol mixture for 2–3 h followed by 45% acetic acid treatment for 30 min. Suitable anthers were smeared with 1% aceto-carmine stain and metaphase I stages of wild type male and GyM-H flowers were documented. Photomicrographs were taken with Nikon Eclipse E200 Microscope using Sony Cybershot DSC-

W320 Camera (digitalized with Optical Zoom - ×4, 14.1 megapixels) with ×10 eye-piece and ×100 oil immersion lens. Each photograph was suitably enlarged and digitally processed under horizontal and vertical resolution at 350 dpi for male and female mitotic metaphase chromosomes and at 72 dpi for GyM mitotic chromosomes. Meiotic metaphase chromosomes of male and GyM-H flower bud were also processed at 350 dpi for better resolution.

Mating design and fruit set analysis

To test the fertility of pollens from male and GyM plants of *C. grandis*, four controlled cross experiments and one self-pollination experiment were designed (Table 2). GyM-H flowers were emasculated before mating. Ten flowers were bagged for each experimental set. The cotton cloth bag was removed after 7 days of controlled pollination and observation was made to each of the flowers. All experimental sets were repeated thrice.

Pollen germination and viability test

To determine the germination rate, fresh pollens from mature male and GyM-H flowers were incubated in germination media with different concentrations of sucrose (1%, 2%, 5%, 10% and 20%) containing 2 mM Ca(NO$_3$)$_2$ and 2 mM H$_3$BO$_3$ [60]. Germination was scored after 1 h of incubation at room temperature. In order to check the fertility of pollens from mature flowers of male plant, GyM-H and converted flowers of AgNO$_3$ treated female plant; pollen grains of each kind were further stained with 1% aceto-carmine solution for 5 min and were documented under the light microscope. Fluorescein diacetate (FDA) test was performed to check the viability of pollens according to the protocol as described by Heslop-Harrison and Heslop-Harrison [61].

Identification and isolation of full length *AGAMOUS* (*CgAG*) and *PISTILLATA* (*CgPI*) homologs

To isolate AGAMOUS (*CgAG*) and PISTILLATA (*CgPI*) homologs from *C. grandis*, total RNA was isolated from harvested flower buds and pooled from all three sexual forms. Degenerate primers (Additional file 11: Table S3) were designed from conserved sequences of PI and AG homologs (Additional file 12: Table S4) using iCODE-HOP [62]. Approximately, 2 μg of total RNA was used for RT-PCR reactions using SuperScript® III One-Step RT-PCR System with Platinum®Taq (Invitrogen-12574-018). The first step of reaction included incubation at 50°C for 20 min for cDNA synthesis followed by 94°C for 2 min, 40 cycles of incubations at 94°C for 15 s, 50°C for 30 s and 68°C for 35 s. Final extension was carried out at 68°C for 5 min. Amplified products were resolved on 2% agarose gel and cloned into pGEMT vector and finally sequence verified. These sequences were used to design primers for 5′ and 3′ RACE to obtain the full

length transcript sequences (Additional file 11: Table S3). RACE ready cDNAs were generated using SMARTer RACE cDNA synthesis kit (Clontech). 5′ and 3′ sequences were further amplified from the cDNAs using the designed primers and the universal primer provided with the kit. Amplified 5′ and 3′ regions of *CgPI* and *CgAG* were sequence verified. Primers were designed to amplify full length transcripts. Deduced amino acid sequences were aligned with other PISTILLATA and AGAMOUS like genes using Clustal Omega and consensus sequences were shaded using Boxshade server [55]. Conserved domains were identified using NCBI's Conserved Domain Database (CDD) search [63].

Quantitative real time PCR (qRT-PCR) analysis

For expression analysis, qRT-PCR was carried out using RNA extracted from whole flower buds at each of the three stages (early, middle and late). RNA was also extracted from the stamens of flowers (late developmental stage) of male, female, GyM-H, GyM-F and converted flowers of AgNO$_3$ treated female plant using the RNeasy Plant Mini Kit (Qiagen) as per the manufacturer's instructions. The yield and RNA purity was determined using Nanodrop 2000c Spectrophotometer (Thermo Scientific, Wilmington, USA) and visualized by gel electrophoresis. Two hundred nanograms (200 ng) of total RNA was used for complementary DNA (cDNA) synthesis by SuperScript III reverse transcriptase (Invitrogen) using an oligo(dT) primer for *CgPI* and *CgAG* genes. 18S rRNA gene was used for normalization for all the reactions. For 18S, fifty nanograms (50 ng) of total RNA was used for complementary DNA (cDNA) synthesis using gene specific reverse primer (Additional file 11: Table S3). qRT-PCR was performed on Roche LightCycler 96 with gene specific forward and reverse primers (Additional file 11: Table S3). The reactions were carried out using KAPA SYBR green master mix (Kapa Biosystems) and incubated at 95°C for 5 min followed by 40 cycles of 95°C for 10 s and 60°C for 20 s. PCR specificity was checked by melting curve analysis, and data were analyzed using the $2^{-\Delta\Delta CT}$ method [64].

Foliar spray of AgNO$_3$ in female and GyM plants

In order to assess the AgNO$_3$ effect, different concentrations of AgNO$_3$ solutions (20 mM, 25 mM, 30 mM, 35 mM and 40 mM) were periodically sprayed on the leaves of female and GyM plants (Additional file 8: Table S2) prior to flowering stage. After 12 days of foliar spray of 35 mM AgNO$_3$ solution, the converted flower buds were harvested at different stages and fixed in 1:3 acetic acid-ethanol mixture for stage specific histological study.

All the supporting data are included as additional files only.

Additional files

Additional file 1: Figure S1. Floral phenotypes on gynomonoecious (GyM) plant. GyM plant showing both hermaphrodite (GyM–H) and pistillate (GyM–F) flowers on the same twig.

Additional file 2: Figure S2. Morphology of hermaphrodite (GyM–H) flowers of gynomonoecious (GyM) plant. Mature hermaphrodite (GyM-H) flower of gynomonoecious (GyM) showing incomplete development of stamens (A) and petaloid stamens (C). Longitudinal sections of early developmental stage of hermaphrodite (GyM–H) flower of gynomonoecious (GyM) plant (B). p: petals, s: sepals, c: carpels, st: stamens, rst: rudimentary stamens, o: ovary, pst: petaloid stamens. Scale bars are 1 cm in A and 1 mm in B.

Additional file 3: Figure S3. Neighbor-joining phylogeny for *Coccinia* based on plastid DNA sequences. The concatenated *mat*K (605 bp) and *trn*S^{GCU}-*trn*G^{UCC} intergenic spacer region (689 bp) were developed by editing and aligning the mentioned species sequences (Additional file 4: Table S1) in MEGA 5.1 software (Tamura K *et al.*, [65]) with *Cucumis sativus* as an outgroup. The numbers at the branches of the tree represent the bootstrap support from 500 replicates. Species names follow Holstein and Renner [37] except for the gynomonoecious (GyM) sexual form of *Coccinia grandis*.

Additional file 4: Table S1. List of accession numbers of the sequences of the species used in phylogenetic tree for both *mat*K and *trn*S^{GCU}-*trn*G^{UCC} intergenic spacer. All other sequences are used from previous work of Holstein and Renner [37].

Additional file 5: Figure S4. Analysis of seed content in *Coccinia grandis*. The seeds from the respective fruits were washed, counted and weighed for evaluating seed production per fruit. (A) Graphical representation of weight of seeds per fruit of female and gynomonoecious (GyM) plants. In the graph, the means ± s.e. are reported (*P <0.05, *t*-test); n = 10. (B) Graphical representation of average number of seeds per fruit of female and gynomonoecious (GyM) plants. In the graph, the means ± s.e. are reported (*P <0.05, *t*-test); n = 10.

Additional file 6: Figure S5. Flower development in *Coccinia grandis*. Developmental stages of the flowers are assigned according to the length of the flower buds. (A) Male, (B) female and (C) gynomonoecious (GyM) flower buds. Scale bars =1 cm.

Additional file 7: Figure S6. Longitudinal sections (L.S) of staminate flower buds of male plant showing pollen development. (A) and (B) are the sections of staminate flower of stages 8 and 12 respectively. p: Petals, st: stamens, pg: pollen grains. Scale bars are 2 mm.

Additional file 8: Table S2. Sex modification in pistillate flower of *Coccinia grandis* female plant after treatment with different doses of silver nitrate.

Additional file 9: Figure S7. Effects of silver nitrate (AgNO$_3$) solution on flower development of gynomonoecious (GyM) plant. (A-D) Longitudinal sections of flowers at different developmental stages from silver nitrate treated gynomonoecious (GyM) plant (after spraying of 35 mM silver nitrate solution). p, petals; s, sepals; c, carpels; st, stamens; rst, rudimentary stamens; o, ovary. Scale bars are 1 cm in A; 2 mm in B,C and D.

Additional file 10: Figure S8. Gynomonoecious *Coccinia grandis* with female and hermaphrodite flowers. (Herbarium Voucher: TU Campus, Karmakar, 433).

Additional file 11: Table S3. List of primers.

Additional file 12: Table S4. List of accession numbers of the sequences of the species used for designing degenerate primers.

Author details
[1]Indian Institute of Science Education and Research (IISER Pune), 900 NCL Innovation Park, Dr. Homi Bhabha road, Pune 411 008, Maharashtra, India. [2]Department of Botany, Tripura University, Suryamaninagar, Tripura 799 022, India.

References
1. Yampolsky C, Yampolsky H: **Distribution of sex forms in the phanerogamic flora.** *Bibliogr Genet* 1922, **3**:1–62.
2. Renner SS, Ricklefs RE: **Dioecy and its correlates in the flowering plants.** *Am J Bot* 1995, **82**(5):596–606.
3. Charlesworth D: **Plant sex determination and sex chromosomes.** *Heredity* 2002, **88**:94–101.
4. Ainsworth C, Parker J, Buchanan-Wollaston V: **Sex determination in plants.** *Curr Top Dev Biol* 1998, **38**:167–223.
5. Guttman DS, Charlesworth D: **An X-linked gene with a degenerate Y-linked homologue in a dioecious plant.** *Nature* 1998, **393**(6682):263–266.
6. Ainsworth C: **Boys and girls come out to play: the molecular biology of dioecious plants.** *Ann Bot* 2000, **86**(2):211–221.
7. Moore RC, Kozyreva O, Lebel-Hardenack S, Siroky J, Hobza R, Vyskot B, Grant SR: **Genetic and functional analysis of DD44, a sex-linked gene from the dioecious plant *Silene latifolia*, provides clues to early events in sex chromosome evolution.** *Genetics* 2003, **163**(1):321–334.
8. Rautenberg A, Sloan DB, Aldén V, Oxelman B: **Phylogenetic relationships of *Silene multinervia* and *Silene* section *Conoimorpha* (Caryophyllaceae).** *Syst Bot* 2012, **37**(1):226–237.
9. Dellaporta SL, Calderon-Urrea A: **Sex determination in flowering plants.** *Plant Cell* 1993, **5**(10):1241–1251.
10. Ming R, Bendahmane A, Renner SS: **Sex chromosomes in land plants.** *Annu Rev Plant Biol* 2011, **62**:485–514.
11. Westergaard M: **Studies on cytology and sex determination in polyploid forms of *Melandrium album*.** *Dansk Botanisk Arkiv* 1940, **10**:1–131.
12. Matsunaga S, Hizume M, Kawano S, Kuroiwa T: **Cytological analyses in *Melandrium album*: genome size, chromosome size and fluorescence *in situ* hybridization.** *Cytologia* 1994, **59**(1):135–141.
13. Westergaard M: **The mechanism of sex determination in dioecious flowering plants.** *Adv Genet* 1958, **9**:217–281.
14. Farbos I, Veuskens J, Vyskot B, Oliveira M, Hinnisdaels S, Aghmir A, Mouras A, Negrutiu I: **Sexual dimorphism in white campion: deletion on the Y chromosome results in a floral asexual phenotype.** *Genetics* 1999, **151**(3):1187–1196.
15. Parker JS, Clark MS: **Dosage sex-chromosome systems in plants.** *Plant Sci* 1991, **80**(1–2):79–92.
16. Ainsworth C, Crossley S, Buchanan-Wollaston V, Thangavelu M, Parker J: **Male and female flowers of the dioecious plant sorrel show different patterns of MADS box gene expression.** *Plant Cell* 1995, **7**(10):1583–1598.
17. Yin T, Difazio SP, Gunter LE, Zhang X, Sewell MM, Woolbright SA, Allan GJ, Kelleher CT, Douglas CJ, Wang M, Tuskan GA: **Genome structure and emerging evidence of an incipient sex chromosome in *Populus*.** *Genome Res* 2008, **18**(3):422–430.
18. Yu Q, Navajas-Pérez R, Tong E, Robertson J, Moore P, Paterson A, Ming R: **Recent origin of dioecious and gynodioecious Y chromosomes in papaya.** *Tropical Plant Biol* 2008, **1**(1):49–57.
19. Zhang J, Boualem A, Bendahmane A, Ming R: **Genomics of sex determination.** *Curr Opin Plant Biol* 2014, **18**:110–116.
20. Bergero R, Charlesworth D: **Preservation of the Y transcriptome in a 10-million-year-old plant sex chromosome system.** *Curr Biol* 2011, **21**(17):1470–1474.
21. Chibalina Margarita V, Filatov Dmitry A: **Plant Y chromosome degeneration is retarded by haploid purifying selection.** *Curr Biol* 2011, **21**(17):1475–1479.
22. Muyle A, Zemp N, Deschamps C, Mousset S, Widmer A, Marais GAB: **Rapid *de novo* evolution of X chromosome dosage compensation in *Silene latifolia*, a plant with young sex chromosomes.** *PLoS Biol* 2012, **10**(4):e1001308.
23. Na J-K, Wang J, Murray J, Gschwend A, Zhang W, Yu Q, Perez R, Feltus F, Chen C, Kubat Z, Moore P, Jiang J, Paterson A, Ming R: **Construction of physical maps for the sex-specific regions of papaya sex chromosomes.** *BMC Genomics* 2012, **13**(1):176.
24. Wang J, Na J-K, Yu Q, Gschwend AR, Han J, Zeng F, Aryal R, VanBuren R, Murray JE, Zhang W, Navajas-Pérez R, Feltus FA, Lemke C, Tong EJ, Chen C, Man Wai C, Singh R, Wang M-L, Min XJ, Alam M, Charlesworth D, Moore PH, Jiang J, Paterson AH, Ming R: **Sequencing papaya X and Y^h chromosomes reveals molecular basis of incipient sex chromosome evolution.** *Proc Natl Acad Sci* 2012, **109**(34):13710–13715.
25. Gschwend AR, Yu Q, Tong EJ, Zeng F, Han J, VanBuren R, Aryal R, Charlesworth D, Moore PH, Paterson AH, Ming R: **Rapid divergence and expansion of the X chromosome in papaya.** *Proc Natl Acad Sci* 2012, **109**(34):13716–13721.
26. Chailakhyan MK: **Genetic and hormonal regulation of growth, flowering, and sex expression in plants.** *Am J Bot* 1979, **66**(6):717–736.
27. Lazarte JE, Garrison SA: **Sex modification in *Asparagus officinalis* L.** *J Am Soc Hortic Sci* 1980, **105**:691–694.
28. Law TF, Lebel-Hardenack S, Grant SR: **Silver enhances stamen development in female white campion (*Silene latifolia* [Caryophyllaceae]).** *Am J Bot* 2002, **89**(6):1014–1020.
29. Stankovic L, Prodanovic S: **Silver nitrate effects on sex expression in cucumber.** *Acta Horticult* 2002, **579**:203–206.
30. Beyer EM: **Silver ion: a potent anti-ethylene agent in cucumber and tomato.** *HortSci* 1976, **11**:175–196.
31. Zhao XC, Qu X, Mathews DE, Schaller GE: **Effect of ethylene pathway mutations upon expression of the ethylene receptor ETR1 from *Arabidopsis*.** *Plant Physiol* 2002, **130**(4):1983–1991.
32. Thomas DT: ***In vitro* modification of sex expression in mulberry (*Morus alba*) by ethrel and silver nitrate.** *Plant Cell Tissue Organ Culture* 2004, **77**(3):277–281.
33. Janoušek B, Široký J, Vyskot B: **Epigenetic control of sexual phenotype in a dioecious plant *Melandrium album*.** *Mol Gen Genet MGG* 1996, **250**(4):483–490.
34. Uchida W, Matsunaga S, Sugiyama R, Kazama Y, Kawano S: **Morphological development of anthers induced by the dimorphic smut fungus *Microbotryum violaceum* in female flowers of the dioecious plant *Silene latifolia*.** *Planta* 2003, **218**(2):240–248.
35. Pfent C, Pobursky KJ, Sather DN, Golenberg EM: **Characterization of *SpAPETALA3* and *SpPISTILLATA*, B class floral identity genes in *Spinacia oleracea*, and their relationship to sexual dimorphism.** *Dev Genes Evol* 2005, **215**(3):132–142.
36. Urasaki N, Tarora K, Shudo A, Ueno H, Tamaki M, Miyagi N, Adaniya S, Matsumura H: **Digital transcriptome analysis of putative sex-determination genes in papaya (*Carica papaya*).** *PLoS One* 2012, **7**(7):e40904.
37. Holstein N, Renner SS: **A dated phylogeny and collection records reveal repeated biome shifts in the African genus *Coccinia* (Cucurbitaceae).** *BMC Evol Biol* 2011, **11**:28.
38. Kumar LSS, Viseveshwaraiah S: **Sex mechanism in *Coccinia indica* Wight and Arn.** *Nature* 1952, **170**(4321):330–331.
39. Chakravorti AK: **Cytology of *Coccinia Indica* W. & A. with reference to the behaviour of its sex-chromosomes.** *Proc Indian Acad Sci B* 1948, **27**(3):74–86.
40. Bhaduri PN, Bose PC: **Cyto-genetical investigations in some common cucurbits, with special reference to fragmentation of chromosomes as a physical basis of speciation.** *J Genet* 1947, **48**(2):237–256.
41. Roy RP, Roy PM: **Mechanism of sex determination in *Coccinia indica*.** *J Indian Bot Soc* 1971, **50A**:391–400.
42. Holstein N: **Evolution, biogeography, and monographic treatment of *Coccinia* (Cucurbitaceae).** In *PhD thesis*. Munich: Ludwig-Maximilians University (LMU); 2012.

43. Sousa A, Fuchs J, Renner SS: **Molecular cytogenetics (FISH, GISH) of** *Coccinia grandis* : a ca. 3 myr-old species of cucurbitaceae with the largest Y/autosome divergence in flowering plants. *Cytogenet Genome Res* 2013, **139**(2):107–118.

44. Ye D, Oliveira M, Veuskens J, Wu Y, Installe P, Hinnisdaels S, Truong AT, Brown S, Mouras A, Negrutiu I: **Sex determination in the dioecious** *Melandrium*. **The X/Y chromosome system allows complementary cloning strategies.** *Plant Sci* 1991, **80**(1–2):93–106.

45. Grant S, Houben A, Vyskot B, Siroky J, Pan W-H, Macas J, Saedler H: **Genetics of sex determination in flowering plants.** *Dev Genet* 1994, **15**(3):214–230.

46. Parker JS: **Sex chromosomes and sexual differentiation in flowering plants.** *Chromosomes Today* 1989, **10**:87–98.

47. Guha A, Sinha RK, Sinha S: **Average packing ratio as a parameter for analyzing the karyotypes of dioecious cucurbits.** *Caryologia* 2004, **57**(1):117–120.

48. Bhowmick BK, Jha TB, Jha S: **Chromosome analysis in the dioecious cucurbit** *Coccinia grandis* (L.) **Voigt.** *Chromosome Sci* 2012, **15**(1):9–15.

49. Lardon A, Georgiev S, Aghmir A, Le Merrer G, Negrutiu I: **Sexual dimorphism in white campion: complex control of carpel number is revealed by Y chromosome deletions.** *Genetics* 1999, **151**(3):1173–1185.

50. Ming R, Wang J, Moore PH, Paterson AH: **Sex chromosomes in flowering plants.** *Am J Bot* 2007, **94**(2):141–150.

51. Scutt CP, Li T, Robertson SE, Willis ME, Gilmartin PM: **Sex determination in dioecious** *Silene latifolia*. **Effects of the Y chromosome and the parasitic smut fungus (***Ustilago violacea***) on gene expression during flower development.** *Plant Physiol* 1997, **114**(3):969–979.

52. Hardenack S, Ye D, Saedler H, Grant S: **Comparison of MADS box gene expression in developing male and female flowers of the dioecious plant white campion.** *Plant Cell* 1994, **6**(12):1775–1787.

53. Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, Sari M-A, Collin F, Flowers JM, Pitrat M, Purugganan MD, Dogimont C, Bendahmane A: **A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons.** *Science* 2008, **321**(5890):836–838.

54. Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C, Bendahmane A: **A transposon-induced epigenetic change leads to sex determination in melon.** *Nature* 2009, **461**(7267):1135–1138.

55. Boualem A, Troadec C, Kovalski I, Sari M-A, Perl-Treves R, Bendahmane A: **A conserved ethylene biosynthesis enzyme leads to andromonoecy in two** *Cucumis* **species.** *PLoS One* 2009, **4**(7):e6144.

56. Lafuma L, Maurice S: **Reproductive characters in a gynodioecious species,** *Silene italica* (Caryophyllaceae), with attention to the gynomonoecious phenotype. *Biol J Linn Soc* 2006, **87**(4):583–591.

57. Dufay M, Lahiani E, Brachi B: **Gender variation and inbreeding depression in gynodioecious-gynomonoecious** *Silene nutans* (Caryophyllaceae). *Int J Plant Sci* 2010, **171**(1):53–62.

58. Cai S, Lashbrook CC: **Laser capture microdissection of plant cells from tape-transferred paraffin sections promotes recovery of structurally intact RNA for global gene profiling.** *Plant J* 2006, **48**(4):628–637.

59. Sharma AK, Sharma A: *Chromosome Techniques: Theory and Practice.* London, UK: Butterworths Co. Ltd; 1980.

60. Brewbaker JL, Kwack BH: **The essential role of calcium ion in pollen germination and pollen tube growth.** *Am J Bot* 1963, **50**(9):859–865.

61. Heslop-Harrison J, Heslop-Harrison Y: **Evaluation of pollen viability by enzymatically induced fluorescence; intracellular hydrolysis of fluorescein diacetate.** *Stain Technol* 1970, **45**(3):115–120.

62. Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S: **Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences.** *Nucleic Acids Res* 1998, **26**(7):1628–1635.

63. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH: **CDD: a conserved domain database for the functional annotation of proteins.** *Nucleic Acids Res* 2011, **39**(suppl 1):D225–D229.

64. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the** $2^{-\Delta\Delta CT}$ **method.** *Methods* 2001, **25**(4):402–408.

65. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: **MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Mol Biol Evol* 2011, **28**(10):2731–2739.

CURRICULUM VITAE

DEVANI RAVI S.

Biology Division,
Indian Institute of Science Education and Research (IISER Pune),
Dr. Homi Bhabha road, Pashan,
Pune-411008, Maharashtra, India
Phone: +91-7276665612, +91-9119441632
E-mail: ravi.devani@students.iiserpune.ac.in,
        ravi.devani89@gmail.com

Permanent Address

c/o Mr. S.K. Devani,
Garden Hill Appartment,
Near Lakshminarayan temple,
Raijibaug-2, Junagadh, Gujarat, India.

Educational Qualifications--

Examination Passed		Year	Board/ University	School / College	Percent total Marks/ CGPA*
M.Sc. Biotechnology	Sem IV	2012	**University of Pune, Pune**	**Department of Biotechnology, University of Pune**	CGPA= 5.38 Out of 6
	Sem III	2011	-do-	-do-	CGPA=4.75 Out of 6
	Sem II	2011	-do-	-do-	CGPA=5.11 Out of 6
	Sem I	2010	-do-	-do-	CGPA=4.87 Out of 6
	M.Sc. CGPA = 5.04/6.0 **"O Grade"**				

B.Sc. Biotechnology	IIIrd Yr	2010	**Saurastra University, Rajkot**	**Shree M & N Virani Science College, Rajkot**	76.25%
	IInd Yr	2009	-do-	-do-	74.5%
	Ist Yr	2008	-do-	-do-	80.5%
	B.Sc. percentage = 77%				
12ᵗʰ		2007	**Gujarat Board, Gandhinagar**	Shree Vivekanand Vinay Mandir, Junagadh.	65.6%
10ᵗʰ		2005	**Gujarat Board, Gandhinagar**	Indian Rayon (Eng.Med.) School, Veraval.	79.57%

Awards and Scholarships--

- **CSIR-JRF- ALL INDIA RANK "2" (Jun'12)**

- **SHYAMA PRASAD MUKHERJEE (SPM) FELLOW**

- **CSIR-JRF- ALL INDIA RANK "45" (Dec'12)**

- **CSIR-JRF- ALL INDIA RANK "45" (Dec'11)**

- **DBT-JRF (2012)**

- **GATE 2012- ALL INDIA RANK "43" (GATE SCORE-607)**

- **GATE 2011- ALL INDIA RANK "260" (GATE SCORE-510)**

- Selected in M.Sc. Biotechnology at PUNE UNIVERSITY, PUNE through **All India Combined Biotechnology Entrance Examination (JNU-CEEB) conducted by Jawaharlal Nehru University**, **New Delhi.** India and receiving a **monthly scholarship** by **Department of Biotechnology**, **Govt. of India**, for two years during M.Sc. Session 2010-2012.

Publications--

- Ghadge A., Karmakar K., **Devani R.S.**, Banerjee J., Mohanasundaram B., Sinha R.K., Sinha S., Banerjee A.K. (2014). Flower development, pollen fertility and sex expression analyses of three sexual phenotypes of *Coccinia grandis*. **BMC Plant Biology.** 14: 325. DOI: 10.1186/s12870-014-0325-0

- **Devani R.S.**, Sinha S., Banerjee J., Sinha R.K., Bendahmane A. and Banerjee A.K. (2017). *De novo* transcriptome assembly from flower buds of dioecious, gynomonoecious and chemically masculinized female *Coccinia grandis* reveals genes associated with sex expression and modification. **BMC Plant Biology.** DOI:10.1186/s12870-017-1187-z

- **Devani R.S.**, Chrimade T., Sinha S., Bendahmane A., Banerjee A.K.* and Banerjee J*. (2018). Flower bud proteome reveals modulation of sex-biased proteins potentially associated with sex expression and modification in dioecious *Coccinia grandis*. **Under review.**

- Sinha S., Karmakar K., **Devani R.S.**, Banerjee J., Sinha R.K., Banerjee A.K. (2016). Preparation of Mitotic and Meiotic Metaphase Chromosomes from Young Leaves and Flower Buds of *Coccinia grandis*. **bio-protocol**. 6(7), e1771. DOI: https://doi.org/10.21769/BioProtoc.1771

- Kondhare K. R., Malankar N., **Devani R. S.** and Banerjee A. K. (2018) Genome-wide transcriptome analysis reveals small RNA profiles involved in early stages of stolon-to-tuber transitions in potato under photoperiodic conditions. *Under Review*.

- **Devani R.S.** *et al.* (2018). A virus induced gene silencing tool for the study of gene function in dioecious *Coccinia grandis*. **Under preparation.**

Workshop and seminars--

- Delivered oral presentation on **"An integrated transcriptomic and proteomic approach to understand sex expression and modification in *Coccinia grandis* - a dioecious cucurbit"** in International Conference on Plant Developmental Biology (ICPDB-2017) held at NISER, Bhubaneswar.
- Presented poster on **"An omics approach to understand sex expression and modification in *Coccinia grandis* - a dioecious member of Cucurbitaceae"** in Indian Society for Developmental biology meeting (InSDB-2017) held at IISER, Pune.
- Presented poster on **"Understanding sex expression and modification in Coccinia grandis - a dioecious member of Cucurbitaceae"** in 3rd International Plant Physiology Congress (IPPC-2015) held at JNU, New Delhi.
- Participated in organization and developed webpage for 35th Annual Plant Tissue Culture Association meeting -2014 held at IISER, Pune.

- Delivered oral presentation on **"Characterization of sexual phenotypes of *Coccinia grandis* (L.): A potential model system to understand dioecy"** in Indian Society for Developmental biology meeting (InSDB-2013) held at TIFR, Mumbai.
- M.Sc. dissertation project on **"Bioaccumulation of chromium in plants: a possible route for chromium oxide nanoparticle synthesis"** under the guidance of Dr. Sujatha Raman, Department of Biotechnology, and Dr. Suresh Gosavi, Department of Physics, University of Pune.
- Certificate course on **Career Oriented Programme (COP)** in **"Plant Tissue Culture"** conducted by Saurashtra University consisting of 6 months training in theoretical and practical aspects.
- Summer project at IISER, Pune under guidance of Dr. Farhat Habib on "**Configuration of GBrowse-genome Browser to display sequence data of *Hydra magnipapillata***" during June-July'11.
- Poster presentation on **"Novel Formulations for Drug Delivery"** in state level SCIFEST event conducted by Shree M.&N. Virani Science College, Rajkot.
- Actively Participation and in an organizing Committee in "**National level seminar on Nanobiotechnology"** jointly organize by Shree M & N Virani Science College, Rajkot and UGC in December 2009.
- Attended a National Level Seminar on "**Biotechnology led paradigm shift-2010"** organize by ARIBAS, V.V. Nagar and UGC in January 2010.
- Actively participated in **Crash Workshop in Biotechnology** organized by Junagadh Agriculture university and Gujarat State Biotechnology Mission in May 2010.

Computer Experience--

- Elementary knowledge of Linux, C language, MS word, MS Excel, MS Power Point and Internet.
- Knowledge of software used for NGS data analysis and proteomic data analysis

Area of Interest

- Plant Molecular Biology, Plant Biotechnology, Bioinformatics

Personal Profile

- Father's name: **Mr. S.K. DEVANI**.
- Mother's name: **Mrs. SHILA DEVANI**.
- Date of Birth: **26-11-1989**
- Languages known: **English, Hindi & Gujarati**