Identification of Candidate Genes, Governing Gynoecy in Bitter Gourd (*Momordica Charantia* L.) by *In-Silico* Gene Expression Analysis

Raghavendra Gunnaiah[1], Vinod M. S.[2], Krishna Prasad[3], Elangovan M[4]*
*Corresponding Author

[1,2,3,4]*Biotech R & D center, Advanta Limited, Lab No. 6, IKP Knowledge Park, Genome Valley, Turkapally, Shameerpet, Hyderabad, India 500078*

ABSTRACT
Bitter gourd (*Momordica charantia* L.) is both medicinally and economically highly valued cucurbit. Hybrid vigor has been well exploited for early maturity, higher yields and other genetic and agronomic traits. However, hybrid seed production in bitter gourd is labor-intensive requiring manual bagging and hand pollination. Use of gynoecious lines as female parent that produce only female flowers have not only reduced the labor requirement, but also have increased yields in hybrids and have good combining ability. However, early phenotypic detection of gynoecy in breeding lines is difficult as gynoecy is highly influenced by environment. To identify the candidate genes governing gynoecy, differentially expressing gene between a gynoecious (Gy323) and a monoecious line (DRAR-1) were explored. 7865 genes were differentially regulated between two lines and 4131 genes were up regulated in gynoecious lines. Up regulated genes in gynoecious line; BTB/POZ domain-containing protein that regulate floral development and previously reported ethylene biosynthesis and regulation genes in cucumber and melon; 1-aminocyclopropane-1-carboxylate oxidases and 1-aminocyclopropane-1-carboxylate synthase and ethylene receptor and ethylene responsive proteins were identified as putative candidate genes for gynoecy in bitter gourd.

Keywords
Bitter gourd, gynoecy, in-silico gene expression analysis, ethylene biosynthesis, BTB/POZ domain-containing protein, ACS1, ACO

1. Introduction
Bitter gourd or bitter melon (*Momordica charantia L.)* is an economically important cucurbitaceous vegetable in terms of both culinary and medicinal usage. Most widely consumed immature fruit is a good source of vitamin C, and also contains vitamin A, phosphorus, and iron. Several bioactive compounds belonging to lectins, ribosome-inactivating proteins, ribonucleases, protease inhibitors, and saponins with potential medicinal values have been isolated and tested for their health benefits [1]. It has been reported to possess antitumor, anti-HIV, anti-diabetic, anti-obesity, and pesticidal activities.

Heterosis is well exploited in bitter gourd for early maturity, higher yield and other agronomic traits through development of hybrids. Heterosis in bitter gourd for yield per vine ranges from 27.3 to 86.1% over the better parent [2,3]. However, hybrid seed production in bitter gourd is labor-intensive requiring manual bagging and hand pollination thereby increasing the cost of seed production. In addition, genetic purity is compromised in the flowers that escape bagging. Use of gynoecious parental lines that produce only pistillate flowers, reduces the labor requirement; increases seed yield and helps to maintain hybrid genetic purity. Gynoecious lines have better genetic combining ability and gynoecious X monoecious hybrids mature early with higher yield potential [4,5]. Like all cucurbits, bitter gourds produce multiple sex forms that are highly influenced by the environment and hormones such as ethylene [6]. Hence, phenotypic identification of pure gynoecious lines is challenging at early stage of plant growth. Identification of gynoecy governing genes and (or) tightly linked markers would ease the identification gynoecious lines and hence their utilization in breeding programme. Gynoecy in bitter gourd was reported to govern by a single recessive gene (*gy-1*) [7]. However, it is also been reported that gynoecious gene might have partially dominant [8] or semi-dominant effect [9]. Nevertheless, molecular mechanism of the gynoecious gene remains un-identified in bitter gourd.

Inheritance and molecular mechanisms of sex expression in cucurbits are well characterized in cucumber (*Cucumis sativus* L.) and melon (*Cucumis melo* L.). Sex expression in cucumber is controlled two loci; dominant F (Femaleness) locus, in combination with another dominant M (monoecious) locus determines the gynoecy [6]. Both the loci play major roles in ethylene biosynthesis. The 'F' locus encodes for an additional copy of 1-aminocyclopropane-1-carboxylic acid synthase (ACS) gene, *CsACS1G*, a key regulator of ethylene biosynthesis [10,11]. The other co-expressing locus 'M', induced by ethylene, further enhances ethylene production by positive feedback regulation [12]. Increased ethylene in the flower primordia arrests the stamens to produce gynoecious flowers.

Similar to cucumber, sex expression in melon is also governed by two major loci; A (andromonoecious) and G (gynoecious) [13,14]. The dominant 'A' allele that suppresses the stamens development in carpel bearing flowers and co-expressing with dominant 'G' allele determine the gynoecy. Further molecular characterizations of 'A' and 'G' loci have established that sex expression in melon is also determined by ethylene. Stamen inhibiting 'A' locus encodes a key ethylene biosynthesis gene *CmACS7* [15]. Other major 'G' locus encodes a WIP transcription factor gene CmWIP1 that was localized to carpel primordia of future stamen flowers. It was reported that expression of *CmWIP1* is antagonistic with *CmACS7*, but doesn't regulate the expression of *CmACS7*.

Molecular characterization of 'F' and 'M' loci in cucumber and 'A' and 'G' loci in melon demonstrate that ethylene plays a significant role in sex determination in cucurbits. In addition to ethylene biosynthesis genes, several ethylene perceptors; *CsETR1*, CsETR2 and *CsERS* gene were also reported to play a significant role in cucurbit sex expression. Taking clues from cucumber and melon, an *in-silico* differential gene expression analysis between a gynoecious and monoecious bitter gourd lines was performed to identify potential candidate genes in ethylene biosynthesis.

2. METHODS

Four sequence read archive (SRA) files, two each from a bitter gourd gynoecious line (Gy323) (SRR947759 and SRR953077) and a monoecious line (DRAR-1) (SRR950973 and SRR953078) were downloaded from NCBI database. The SRA files were generated from the single end reads using Illumina Genome Analyzer, by sequencing the total RNA isolated from pooled bitter gourd tissues (stem, root, leaves and flower buds) (personal communication). SRA data used in this study was deposited by Faculty of Science, Banaras, Hindu University, Lucknow, India on 31-July-2013. Low quality reads and adapter sequences were removed before being used for analyzing the differential gene expression. Since, reference genome sequence of bitter gourd is not available, transcript shotgun assembly (TSA) developed from the de-novo assembly of bitter gourd SRA (NCBI accession numbers GANF01000001-GANF01054667 and GANG01000001-GANG01051326) was used as reference assemblies for aligning short reads.

To develop a common reference assembly, all the assemblies were merged and redundant sequences were removed by using Cd-hit [16] and reassembled with CAP3 [17]. Alignment of short reads to the reference transcript assemblies were carried out using Bowtie 2 [17]. The contig nucleotide assemblies generated from the Bowtie 2 were annotated by BLASTX searching against viridiplantae protein reference sequence database. Gene ontology (GO) terms were assigned to the annotated protein sequences and mapped onto KEGG metabolic pathways using Blast2GO [18].

The fold change expression of differentially expressing genes between gynoecious line and monoecious line was calculated using FPKM (fragments per kilobase of exon per million fragments mapped) as per the following formulae. The differentially expressing genes were mapped on KEGG pathways.

3. RESULTS

3.1 Bitter Gourd Transcriptome

To identify the gynoecy determining candidate genes in bitter gourd, we explored differentially expressed genes between gynoecious (Gy323) and monoecious line (DRAR1). By combining TSA from Gy323 (GANF01000001-GANF01054667) and DRAR1 (GANG01000001-GANG01051326), a common reference transcript assembly was developed. After removing the redundant transcript assemblies, 46782 transcript assemblies with total trancriptome length of 55862771 bp were retained for further annotation and differential gene expression analysis. Contig length of assembled transcripts varied from 200 bp to 11022 bp with an average length of 1194.11 bp and N-50 contig length of 1789 bp. About 47 % of the total 46782 TSA were annotated as protein coding genes and among them 73% of the protein coded genes were annotated with specific biological or molecular function. Majority of protein coding genes were homologous to cucumber proteins and merely 571 and 323 transcripts were homologous to melon and watermelon proteins respectively.

3.2 Differential Gene Expression Between Gynoecious Line And Monoecious Line

Log 2 Fold change (Gy323/DRAR-1) values of +1 and -1 was considered as up regulated and down regulated respectively in the gynoecious line Gy-323. A total of 7865 genes were differentially regulated between two lines, 4131 genes were up-regulated and 3734 genes were down-regulated in gynoecious line Gy-323. Majority of the differentially regulated genes were belonging to metabolic processes (167) and cellular processes (144) (Fig. 1). Two genes with reproductive function, 5 involved

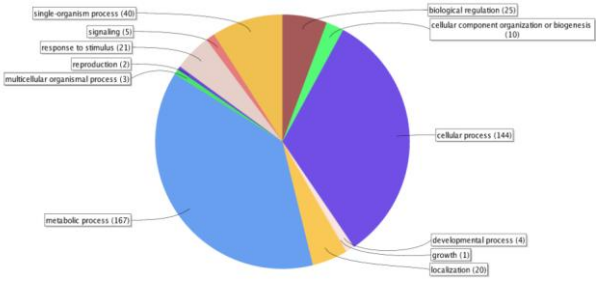

Figure 1 Classification of up regulated genes in gynoecious bitter gourd, according to their biological function

in signaling and 21 that are induced upon external stimulus might play a role in sex expression (Fig. 1). Majority of the genes belonging to these pathways were related to plant hormone signal transduction such as auxin and ethylene biosynthesis. Potential candidate genes of gynoecy that might play a significant role in ethylene biosynthesis or flower organ development, which were up regulated in gynoecious line Gy323, are presented in Table I. Among these candidate genes, a homologue of *Arabidopsis* that determine the floral organ development was up regulated with a higher fold change (6.9). Among the rest, 8 were belonging to ethylene biosynthesis that includes one ACC synthase, 7 ACC oxidases and two ethylene receptors.

4. DISCUSSION

Non-targeted transcript analysis is an efficient tool in identifying candidate genes through differential gene expression analysis. It provides unbiased information on the differential expressed genes between two genes that differ for a specific trait. Differentially expressed genes can be mapped on to metabolic pathways that provide first hand information on the induced metabolic pathways, which can be further explored for mining candidate genes in and around the pathway. In this study, thousands of genes were differentially expressed between gynoecious and monoecious bitter gourd lines that might or might not have any role in governing sex expression in bitter gourd. Such genes mainly belong to hormonal signal transduction and can be further explored for their role in sex expression by loss/gain of gene function studies such as TILLING and virus induced gene silencing. Presence of large number of up regulated genes might be due to pooling of different tissues for RNA extraction in order to enrich the bitter gourd transcriptome. Since, different tissues express multitude of genes, identification of putative candidate genes involved in sex determination is challenging. As an alternative strategy, transcriptome of only flower buds could provide accurate information on gynoecy determining genes in bitter gourd. Nevertheless, the few candidate genes that control sex expression in cucumber, melon and Arabidopsis were up regulated and these genes can be further validated for their role in controlling sex expression in bitter gourd.

4.1 Potential Candidate Genes For Gynoecy In Bitter Gourd

An Arabidopsis homologue BTB/POZ domain-containing protein was up regulated almost seven times in the bitter gourd gynoecious line Gy323. BTB/POZ domain-containing proteins are a family of scaffold proteins, characterized by protein-protein interaction [19]. Few of these auxin responsive proteins (BT1–BT5) perform a crucial role male and female gametophyte development in *Arabidopsis*. Other plant BTB proteins also play

a major role in many cellular processes, such as leaf and flower morphogenesis for BLADE-ON-PETIOLE 1and 2 [20] and in ethylene responses for *ETO1*, *EOL1* and *EOL2* [21]. Role of BTB/POZ domain containing proteins in determining gynoecy in bitter gourd has to be further validated.

Role of ethylene in determining gynoecy in cucurbits is well demonstrated in both cucumber and melon. Lower levels of ethylene are required to stimulate carpel development and further increase in its concentration inhibits the stamen development [12]. Similar to cucumber and melon, genes of ethylene biosynthesis and regulation were up regulated in the bitter gourd gynoecious line, implicating their potential role in determining gynoecy (Fig. 2). An ACS gene, AAQ14268.1, homologous to *CsACS1* along with its recombinant counterpart, branched-chain-amino-acid aminotransferase-like protein 2-like in forming a second copy of *ACS1* gene, *CsACS1G* was also up regulated. *CsACS1G* encodes a major gynoecy determining 'F' locus in

Table 1: Putative candidate genes of ethylene biosynthesis and floral organ development that might determine gynoecy in bitter gourd

NCBI Accession number	Sequence name	Log 2 Fold change	Seq length (bp)	Function
NP_001189642.1	BTB/POZ domain-containing protein [*Arabidopsis thaliana*]	6.92	3091	Proximal/distal pattern formation; flower morphogenesis;
XP_003516383.1	PREDICTED: 1-aminocyclopropane-1-carboxylate oxidase 1-like [*Glycine max*]	3.84	608	Ethylene biosynthesis
XP_004134077.1	PREDICTED: 1-aminocyclopropane-1-carboxylate oxidase homolog 1-like [*Cucumis sativus*]	2.14	1479	Ethylene biosynthesis
O82436.1	ETR1_CUCMR; Ethylene receptor 1 (Cm-ETR1	1.87	3115	Ethylene-activated signaling pathway
XP_004134076.1	PREDICTED: 1-aminocyclopropane-1-carboxylate oxidase homolog 1-like [*Cucumis sativus*]	1.63	1838	Ethylene biosynthesis
XP_004134074.1	PREDICTED: 1-aminocyclopropane-1-carboxylate oxidase homolog 3-like [*Cucumis sativus*]	1.57	776	Ethylene biosynthesis
XP_004491714.1	PREDICTED: 1-aminocyclopropane-1-carboxylate oxidase homolog 1-like [*Cicer arietinum*]	1.54	1607	Ethylene biosynthesis
XP_004134416.1	PREDICTED: 1-aminocyclopropane-1-carboxylate oxidase homolog 1-like [*Cucumis sativus*]	1.49	1296	Ethylene biosynthesis
XP_004142180.1	PREDICTED: 1-aminocyclopropane-1-carboxylate oxidase homolog 6-like isoform 1 [*Cucumis sativus*]	1.46	352	Ethylene biosynthesis
BAD01555.1	ERF-like protein [*Cucumis melo*]	1.44	559	F: sequence-specific DNA binding transcription factor activity;
XP_004142181.1	PREDICTED: 1-aminocyclopropane-1-carboxylate oxidase homolog 6-like isoform 2 [*Cucumis sativus*]	1.440	394	Ethylene biosynthesis
AAQ14268.1	1-aminocyclopropane-1-carboxylate synthase [*Momordica charantia]*	1.409	2199	Ethylene biosynthesis
XP_004139364.1	PREDICTED: branched-chain-amino-acid aminotransferase-like protein 2-like [*Cucumis sativus*]	1.32	2205	Gynoecy determining gene in cucumber

cucumber [10]. It is possible that homolog of *CsACS1G* in bitter gourd might also play a key role in determining gynoecy.

Unlike cucumber and melon, 8 ACC oxidases (ACO) were up regulated in gynoecious bitter gourd. ACO is an ethylene-forming enzyme that catalyzes the final step in ethylene formation, was also implicated in determining different sex genotypes in cucumber [22]. Likewise, an ethylene receptor protein, ETR1_CUCMR and ethylene responsive factor, ERF1 homologs were also up regulated in the gynoecious lines. Cucumber ethylene receptors CsETR1 and CsETR2 have shown higher expression in the shoot apices of gynoecious cucumber [23].

Higher expression of ethylene biosynthesis and regulation genes in the gynoecious line indicates that, similar to cucumber and melon, gynoecy is bitter gourd might also be regulated by ethylene biosynthesis. Putative candidate genes BTB/POZ domain-containing protein, ACS, ACO, ETR and ERFs can be further validated for their expression exclusively in flower primordia and used as potential candidate genes for future breeding program.

5. CONCLUSION

Enormous open access genome sequence and gene expression data is available due to breakthroughs in next generation sequencing technologies. While using various bioinformatics tools, several genes can be annotated and biologically interpreted

using the open access data. Putative ethylene biosynthesis candidate genes identified in this study can be further validated as candidate genes of gynoecy in bitter gourd and exploited in breeding of gynoecious lines for development of hybrids.

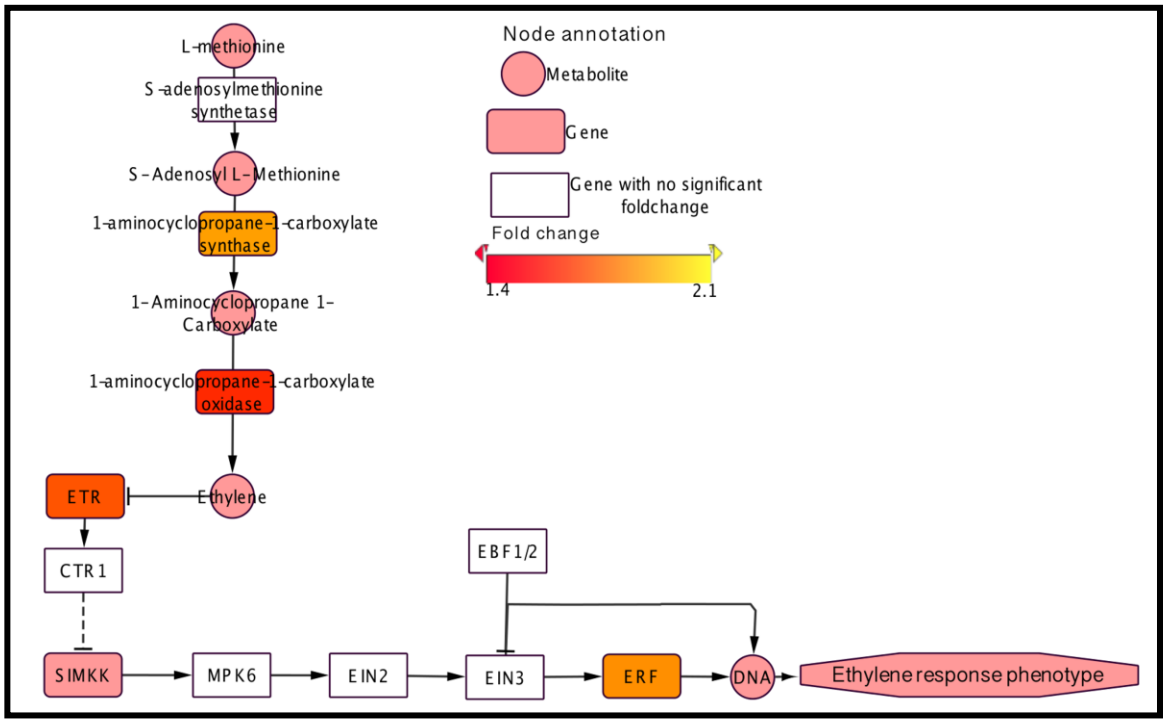

Fig. 02: Ethylene biosynthesis and regulation pathway in gynoecious bitter gourd

References

[1] Fang EF, Ng TB (2013) The Bitter Fruit with Sweet Health Benefits: A Comprehensive Synopsis of Recent Research Progress on Medicinal Properties of *Momordica Charantia*. Antitumor Potential and other Emerging Medicinal Properties of Natural Compounds: Springer. pp. 315-334.

[2] Behera T (2008) Heterosis in bittergourd. Journal of New Seeds 6: 217-221.

[3] Sngh A, Pan R, Bhavana P (2013) Heterosis and Combining Ability Analysis in Bittergourd (*Momordica charantia* L.)

[4] Dey S, Behera T, Munshi A, Pal A (2010) Gynoecious inbred with better combining ability improves yield a

[5] nd earliness in bitter gourd (*Momordica charantia* L.). Euphytica 173: 37-47.

[6] Dey S, Behera T, Munshi A, Rakshit S, Bhatia R (2012) Utility of gynoecious sex form in heterosis breeding of bitter gourd and genetics of associated vegetative and flowering traits. Indian J. Hort. 69: 523-529.

[7] Grumet R, Taft J, YiHong W, Behera T, Kole C (2011) Sex expression in cucurbits. Genetics, genomics and breeding in crop plants: 353-375.

[8] Ram D, Kumar S, Singh M, Rai M, Kalloo G (2006) Inheritance of gynoecism in bitter gourd (*Momordica charantia* L.). J. heredity 97: 294-295.

[9] Iwamoto E, Ishida T (2006) Development of gynoecious inbred line in balsam pear (*Momordica charantia* L.). Hort. Res. 5.

[10] Matsumura H, Miyagi N, Taniai N, Fukushima M, Tarora K, et al. (2014) Mapping of the Gynoecy in Bitter Gourd (*Momordica charantia*) Using RAD-Seq Analysis. PLOS ONE 9: e87138.

[11] Trebitsh T, Staub JE, O'Neill SD (1997) Identification of a 1-aminocyclopropane-1-carboxylic acid synthase gene linked to the female (F) locus that enhances female sex expression in cucumber. Plant Physiol. 113: 987-995.

[12] Mibus H, Tatlioglu T (2004) Molecular characterization and isolation of the F/f gene for femaleness in cucumber (*Cucumis sativus* L.). Theo. Appl. Genet. 109: 1669-1676.

[13] Li Z, Wang S, Tao Q, Pan J, Si L, et al. (2012) A putative positive feedback regulation mechanism in CsACS2 expression suggests a modified model for sex determination in cucumber (*Cucumis sativus* L.). J. Exp. Bot. 63: 4475-4484.

[14] Kenigsbuch D, Y C (1990) The Inheritance of Gynoecy in Muskmelon. Genome 33: 317-320.

[15] Roy R, Saran S (1990) Sex expression in the Cucurbitaceae. Biology and utilization of the Cucurbitaceae Cornell University Press, Ithaca, NY: 251-268.

[16] Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, et al. (2008) A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science 321: 836-838.

[17] Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658-1659.

[18] Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res. 9: 868-877.

[19] Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674-3676.

[20] Robert HS, Quint A, Brand D, Vivian- Smith A, Offringa R (2009) BTB AND TAZ DOMAIN scaffold proteins perform a crucial function in Arabidopsis development. Plant J. 58: 109-121.

[21] Ha CM, Jun JH, Nam HG, Fletcher JC (2004) BLADE-ON-PETIOLE1 encodes a BTB/POZ domain protein required for leaf morphogenesis in Arabidopsis thaliana. Plant and Cell Physiol. 45: 1361-1370.

[22] Wang KL-C, Yoshida H, Lurin C, Ecker JR (2004) Regulation of ethylene gas biosynthesis by the Arabidopsis ETO1 protein. Nature 428: 945-950.

[23] Kahana A, Silberstein L, Kessler N, Goldstein RS, Perl-Treves R (1999) Expression of ACC oxidase genes differs among sex genotypes and sex phases in cucumber. Plant Mol. Biol. 41: 517-528.

[24] Yamasaki S, Fujii N, Matsuura S, Mizusawa H, Takahashi H (2001) The M locus and ethylene-controlled sex determination in andromonoecious cucumber plants. Plant Cell Physiol. 42: 608-619.