DEVELOPMENT OF EXPRESSED SEQUENCE TAGS FOR GENE DISCOVERY, MAPPING AND DISTRIBUTION DURING FRUIT RIPENING IN PINEAPPLES

Kumar, S.V., Ong, W.D., Yew, C.W. and Yusuf, N.H.M

Biotechnology Research Institute, Universiti Malaysia Sabah,
Jalan UMS, 88400 Kota Kinabalu, Sabah

(vijay@ums.edu.my)

ABSTRACT

Pineapple (*Ananas comosus* var. *comosus*), is an important tropical non-climacteric fruit with high commercial potential. Commercial success relies heavily on improving fruit quality through the development of new varieties with novel consumer traits including flavor, texture, appearance and sweetness. To increase our understanding of the genetic diversity and gene-based control of these key traits in pineapple, a collection of approximately 30,000 EST contigs were derived from Sanger-sequencing and Solexa paired-end sequencing. All sequences obtained were screened for similarities using BLAST. Functional classification for gene ontology was carried out based on the annotations available in the GenBank for other plant species. Many potential molecular markers were abundant in the pineapple transcripts. EST-SSRs were detected in 3.5% of the contigs generated. PCR primers were able to be designed for only 26.6% of these EST-SSRs. Pineapple microRNAs from fruit and leaf tissues were also identified and characterised. These miRNAs play an important role in regulating gene expression at the post-transcriptional level through complementary binding to specific mRNA targets and subsequently inducing gene silencing either by translational repression or mRNA degradation. This report provides first-hand account of the pineapple transcriptome.

INTRODUCTION

The genetic pathway underlying non-climacteric fruit development is still unknown. These fruits, such as, pineapples, citrus and grapes, do not show an increase in respiration and ethylene production during the onset of fruit ripening [4]. Transcriptome analysis of the pineapple fruit may provide the framework for the elucidation of the gene expression and regulatory roles played by the transcripts during ripening.

Expressed Sequence Tag (EST) libraries and databases are powerful tools for gene discovery, gene mapping, and for the analysis of quantitative traits. ESTs are generated by large-scale sequencing of randomly picked clones from cDNA libraries constructed from mRNA isolated at a particular development stage and/or tissue. With new technologies, instead of a clone-by clone sequencing approach, the emergence of next-generation sequencing provide a better approach, as sequencing by massively parallel sequencing reduces the costs, time, labour, errors associated by clone mishandling, bias with type of vector used in cloning, and provides the opportunity to capture the rare

transcripts [6]. Apart from long mRNA transcripts, small RNAs with regulatory roles such as microRNAs have been shown to directly regulate the development and morphogenesis of embryo, leaves, flower and roots. The role of miRNAs in fruit ripening remains unclear.

This paper reports the analysis of the pineapple fruit transcriptome using data generated from Sanger's and Solexa mRNA sequencing. It provides a framework of transcripts expressed, with special emphasis on transcripts, which play an important role in controlling pineapple fruit quality traits such as sweetness, aroma, and texture. In addition to the ESTs, miRNAs associated with pineapple fruit development, isolated from a library of small RNAs or through orthology with conserved regions in other plant species are described. Quantification of pineapple miRNAs at two different development stages, i.e. pre-ripening and post-ripening was determined by RT-qPCR.

MATERIALS AND METHODS

Plant materials & total RNA

Fully ripen pineapples were obtained from a pineapple farm in Penampang, Sabah. High quality total RNA was extracted using the protocol of Li *et al*. (2006) with some minor modification. Small RNAs were enriched from 50 µg of total RNA by using mirVana microRNA isolation kit (Ambion).

cDNA library construction & DNA sequencing

cDNA was constructed using the Smart cDNA Library Construction Kit (Clontech). The pineapple mRNA was reversed transcribed to generate cDNAs. The cDNAs was then normalised, digested by *Sfi*l, sized fractionated, ligated into λ TripIEX2 vector, transformed into *E. coli* XL1-Blue, and finally sequenced from the 5' terminal ends. Total RNA was also sequenced by pair-end Solexa sequencing. The sequencing service was provided by Ilumina, USA.

Clustering, assembly, functional annotation & EST-SSR

EST clustering was performed to collect overlapping ESTs from the same transcript of a single gene into unique clusters to reduce redundancy. Paired-end sequences were assembled using Velvet *de novo* assembler. Contigs generated from the assembly were screened for similarities with GenBank's non-redundant protein database by the BLASTX algorithm. A match was considered significant when the score was higher than 100 with E-value scores $\leq 10^{-6}$. Type I SSR motifs were identified from the pineapple ESTs using the SynaRex tool at the local server. The ESTs were screened for the presence of dinucleotide, trinucleotide and tetranucleotide motifs.

Isolation of miRNA and its expression in pineapple fruit

The procedure for the isolation of miRNA was according to Lau *et al.* (2001) with modifications. Enriched small RNAs (from total RNAs) were size fractionated on 15%/8M urea polyacrylamide gel. The gel region within the range of 17-24 nucleotides

was excised, purified, attached to linkers, reverse transcribed, cloned and finally sequenced. Thirteen stem-loop primers were used for RT-qPCR to detect miRNA expression in two developmental stages i.e. pre-ripening and post-ripening stages. The primers were based on those described by Li *et al*. (2009) which was originally derived from *Arabidopsis*.

RESULTS AND DISCUSSION

Sequencing and *de novo* assembly

Pineapple fruit cDNA library was constructed and paired-end Solexa mRNA sequencing was performed. The insert size of the library was 200 bp in length with each forward and reverse reads of 75bp. The sequencing was able to generate 4.5 million reads with a total sequence of 675 Mbp. *De novo* assembly of the Solexa sequences using Velvet software generated approximately 30,000 contigs. The sequencing coverage ranged from 1 to 3,290 fold. The average length of the contigs was approximately 1.6kb in length with the longest at 3,637 bp.

Similarity search and high expression transcript assessment

The BLASTX search showed that the ESTs were related to primary metabolisms, amino acid synthesis and processing, membrane and transport, cell division, cytoskeleton, cell wall and metabolism, signal transduction, defense and stress related protein and also secondary metabolism. Polygalacturonase, an enzyme that was responsible in depolymerization of pectin [5], was highly expressed 1,000 fold in fully ripen pineapple fruits. Polygalacturonase plays an important role in fruit softening compared to fasciclin-like arabinogalactan protein, another enzyme responsible for cell wall activity, which accounted for an expression level of 18 fold. Alcohol acyltransferase and alcohol dehydrogenases, both implicated in fruit aroma, were highly expressed 500 fold in fully ripen fruit. Type I SSRs with the minimum repeat motif of eight for dinucleotide, six for trinucleotides and five for tetranucleotides were identified. Approximately, 1005 contigs generated from the *de novo* assembly contained SSRs. Dinucleotide was the most abundant with 50% of the SSR-containing contigs, followed by trinucleotides and tetranucleotide with 46% and 4%, respectively.

MicroRNA expression

Thirteen miRNAs namely miR156, miR157, miR159, miR162, miR164, miR165, miR167, miR168, miR169, miR170, miR171 and miR390 were found to be expressed in pineapple fruit. Two genes miR168 and miR390 were up-regulated at 1.79 fold and 2.64 fold while another one, miR164 was down-regulated down to 0.07 fold. The genes with stage-regulated expression identified here are candidates for further investigation to reveal the function of miRNAs in non-climacteric fruit ripening.

CONCLUSION

Transcriptome analysis of both ESTs and miRNAs provides a global picture of the underlying expression and regulatory genetic pathways that control fruit ripening in the model tropical non-climacteric fruit, pineapple. The detailed genic architecture will provide a means to develop new improved varieties through marker assisted selection or genetic transformation.

ACKNOWLEDGEMENT

This project was funded by MOSTI Science fund Grant SCF0007-BIO-2006.

REFERENCES

1. Lau, N. C., Lim, L. P., Weinstein, E. G. and Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science **294:** 858-862.

2. Li, H., Zhang, Z., Huang, F., Chang, L. & Ma, Y.(2009). MicroRNA expression profiles in conventional and micropopagated strawberry (*Fragaria X ananassa* Duch.). Plant Cell Rep 10.

3. Li, J.H., Tang, C.H., Song, C.Y., Chen, M.J., Feng, Z.Y. and Pan, Y.J. (2006). A simple, rapid and effective method for total RNA extraction from *Lentinula edodes*. *Biotechnol. Letter*. **26:**1193-1197.

4. Moyle, R., Fairbairn, D. J., Ripi, J., Crowe, M. and Botella, J. R. (2005). Developing pineapple fruit has a small transcriptome dominated by metallothionein. *Journal of Experimental Botany* **409:**101-112.

5. Nishiyama, K., Guis, M., Rose, J.K.C., Kubo, Y., Bennett, K.A., Wangjin, L., Kato, K., Ushijima, K., Nakano, R., Inaba, A., Bauzayen, M., Latche, A., Pech, J.C. and Bennett, A.B. (2007). Ethylene regulation of fruit softening and cell wall disassembly in *Charentais* melon. *Journal of Experimental Botany* **58:**1281-1289.

6. Weber, A.P.M., Weber, K.L., Carr, K., Wilkerson, C., Ohlrogge, J.B. (2007). Sampling the *Arabidopsis* transcriptome by massively parallel pyrosequencing. *Plant Physiology* **144:** 32-42.