Transcriptome profiling of genes that govern the pigment systems of *Bombyx mori* variant Dazao

Archana V. Pawar, Aall H. Manjunatha Reddy

Department of Biotechnology, Rashtreeya Vidyalaya College of Engineering, Bengaluru, Karnataka, India. Corresponding author: A. H. Manjunatha Reddy
ahmanjunatha@rvce.edu.in

Abstract. Sericulture industry has turned out to be one of the highest revenue earners mainly through effective research and innovations. The latest fetish of the research community under sericulture is the production of naturally colored silk, as it avoids the traditional dyeing process that is costly and responsible for water pollution. To promote the principle of ecology with economy, scientists are now studying the pigments systems of silkworms for the production of naturally colored silk. In the present study, an attempt was made to understand the genomic intricacies that characterize each pigment of the silkworm, which can be further correlated for the production of colored silk. With respect to the above objective, transcriptome sequence of silkworm variety Dazao (wild-type and colored) archived in sequence read archive (SRA) were studied using the transcriptome profiling technique of RNA-Seq Technology. The study revealed a total of 17,253 expressed transcripts, out of which 14,548 transcripts were annotated with *Bombyx mori* proteins. These transcripts were further studied for pathway analysis and networking of genes to understand their biological significance in pigment production. On the basis of the above methodology and protocol, genes and their associated pathways were discovered for production melanins, ommochrome, sepiaterin and betalain pigments. Also 5 potentially novel transcripts were discovered during the course of NGS analysis whose functionality can be confirmed through wet lab studies.
Key Words: silkworm, RNA-Seq technology, SILKDB, transcripts, sericulture.

Introduction. Sericulture is the art of rearing silkworms to produce silk. This ancient craft today is responsible for an important chunk in the world economy with China, India, Japan, Korea and Brazil being its key players. Just like ancient times, even today silk reigns supreme as an object of desire and is aptly designated as the “Queen of textiles” and “Fashions biggest canvass” (Gangopadhyay 2008). This industry is also a classic example of an age-old craft backed up with new age technology. The ever increasing demand for silk and cut throat competition at the global frontier has made innovations common in this field. The advent of low cost next generation technologies has provided a new perspective towards understanding the biological application of silkworms that has in turn led to mind boggling innovations in this field. Some of the latest include: strong and tough spider silk (Lewis 2006), glow in dark fluorescent silk (Iizuka et al 2013), silk leather, silk surgical guazes and bandages, biological transplants based on silk proteins, silksheet for artificial fruit ripening (Chen et al 2007) etc.

The latest fetish of the research community under sericulture is the production of naturally colored silk. This topic is of current interest in sericulture research because, till date silk fibers produced by *Bombyx mori* were in shades of cream or white. These fibers were later dyed to obtain a desired color tone and provide a better market appeal. This dyeing of silk is an added financial burden to the industry, also the dyes used are non-biodegradable and responsible for environmental pollution. The dyeing process of silk also makes silk handling difficult for the customer. To promote the principle of ecology with economy, scientists are now studying the pigments systems of silkworms for the production of naturally colored silk. This concept is inspired from the existence of more than 500 different varieties of silkworms in nature that are reported to produce naturally colored silk. However, these wild-type silkworms are not harvested for commercial silk

production as many of them require strict growth conditions failing to which the entire crop is lost. These wild-type silkworms are also more prone to infections and require sterile handling. The post-harvest reeling and conditioning of silk, is also an added burden on the industry owners that makes the exploitation of wild-type silkworms limited for commercial silk production. Thus efforts are being made to understand the genomic elements responsible for pigment production in wild-type silkworm and incorporate or express the same in the commercially used domesticated *B. mori.*

Transcriptome is the set of all proteins expressed at a particular life time of an organ in an organism. The study of the transcriptome of any tissue or organ of an organism is of maximal interest, because it provides a complete view of probable proteins produced by the organism without having to resort to expensive proteomic studies. Also current accomplishments of Next Generation Sequencing (NGS) technologies have made them the current favorite of genomic studies. High quality and affordability of current NGS techniques, combined with extensive computational capabilities has provided transcriptomic analyses a prevalent place in biological studies. Thus in our present study we intended to use transcriptomic profiling to allow a deeper and more detailed exploration of the genomic interplay that was responsible for pigment production in the wild type silkworm Dazao. Transcriptome profiling on the lines of RNA-Seq technology was used as a tool to understand the genomic intricacies that characterized each pigment of the silkworm, which can be further correlated for the production of colored silk. Since transcriptome profiling is a fast, simple and cost effective method to understand the functional elements within the genome and discover novel developmental regulatory networks (Randolph 2006). Such an approach was thought to help provide better insights on the pigment production in *B. mori* variant Dazao.

Review of literature. Over the years, *B. mori* has been the subject of interest for many scientific studies and research work. The reasons for this could be its small size, shorter generation time, ease of handling and its position in the animal kingdom as the model organism for the Lepidopteran group. The advent of NGS technologies has positively impacted the research on *B. mori* and has led to newer accomplishments in the field of sericulture research. The study of literature in the below mentioned fields enabled us to outline the research objective of this project, aggregate resource materials for the intended study and upgrade the available annotation data.

Sequencing projects of B. mori. The genome sequence of *B. mori* was analyzed and the genetic elements were identified and named with the prefix “Bm1” by Adams (1986). It was later revealed by Ogura et al (1994), that the *B. mori* genome is interspersed with repetitive long sequences or fragments called L1. These fragments are 5.1 kb in length and contain 15 bp CpG-rich sequence. Also Zhao et al (2013) reported that about 1.4% of the silk genome is made up of segmental duplications which play an important role in immunity, environmental adaptation, detoxification etc and hence are evolutionarily important entities. The whole genome sequencing analysis revealed that *B. mori* has a total length of 514 Mb with gaps and 97% of the genome was arranged into scaffolds. The sequence data clearly pointed out the presence of repeating elements from truncated transposoons and insertions of mitochondrial elements. Data analysis from *B. mori* genome sequencing project by Jin-Shan et al (2005) and Quan et al (2002) revealed the presence of long terminal retrotransposons (LTR) that make up at least 11.8% of the *B. mori* genome. However, even after several sequencing approaches the, *B. mori* genome contained several unassembled regions that hampers functional annotation and analysis, thus Yamamoto et al (2008) used bacterial artificial chromosome sequencing method to create an integrated map which improve the genome coverage to 76%. This understanding of the *B. mori* genome was further elaborated by the International Silkworm Genome Consortium (2008), where they combined the information obtained from whole-genome shotgun, fosmid- and BAC-end sequencing to produce a sequence map where the functionality of about 87% of the scaffolds were assigned to the 28 chromosomes of *B. mori.*

Public availability of B. mori research data. The increase in quality sequence data of *B. mori* also brought about the problem of managing and understanding of sequence information. To make this easier and allow optimum usage of sequence information a data mining tool called KAIKObase was created by Shimomura et al (2009). KAIKObase incorporated within it sequencing, literature, structural and functional data that would enable a prospective researcher a 360 degree view of sequencing data. Similar approach was followed by Duan et al (2010) who created an integrated silk genome database called SILKDB, incorporating silk genome data, assembly sequencing data along with expression data of microarrays and ESTs of *B. mori*. They also provided specialized tools that would enable viewing of chromosome data, browse available annotation data etc. Also Li et al (2012) studied the entire transcriptome profile of *B. mori* at different developmental stages and identified 5,428 of them were novel exons. To allow optimum utilization of the studied transcriptome data by the research community, the novel finds were incorporated into available silkgenome data of SilkDB and released in the form of a secondary database called SILKTransDB. The data obtained in these databases were further consolidated by Suzuki et al (2013) by performing a full scale cDNA sequencing of 14 different tissues of *B. mori*. This data was used to identify 16,823 gene loci which was validated based on GLEAN gene models, published mRNA and orthology analysis of 153 insect species of the lepidopteran group.

Experimental studies conducted to understand pigment production in B. mori. The interest on silkworm pigments was exhibited in the form of study on the identification of pterine pigment with blue fluorescence of *B. mori* and fluorescyanine by Busnel et al (1950). Hirata et al (1950) deduced that xanthopterin was a component of larval skin based on pigment mutant studies. The hunt for newer silkworm pigments continued and involved the identification of fluorescyanine (ichthyopterin) from *B. mori* eggs and carp scales by Hirata & Nawa (1951) and isolation of fluorescent substance from blood and in the malpighian tubules from normal by Jimbo & Nishiyama (1962) and polyhedral disease infected silkworm by Drilhon (1951). The study was continued to identify similar fluorescent peptide in the hemolymph of paralyzed silkworm Drilhon & Vago (1953). The preceding year saw focused research on synthesis of pterins of *B. mori* by Hirata et al (1952) with specific emphasis on the constitution. A newer take on pigment research came, when it was reported that certain strains of *Bacillus thuringiensis* Berliner were responsible for pigment production in silkworms (Shvetsova & Zurabova 1966). This was followed by identification and genetic and biochemical studies of color imparting chromogranules on silkworm larval skin by Sakurai (1968). Later phenoxazine pigment was discovered as a product obtained from metabolism of 3-oxy-kynurenine by Ishiguro et al (1971). Also the transcriptome sequencing of mutant and wild type silkworms by Quan et al (2002) helped better understand the ommochrome production by the expression of Kynurenine 3-monooxygenase. Xanthomatins or red proteins were also isolated and identified from the body fluid like urine of normal and mutant rb-silkworms by Ishiguro & Nagamura (1971). Later studies were conducted to establish light induced melanin pigment production by Kuan et al (1979). This was followed by the isolation and characterization of a reddish brown melanization hormone by Matsumoto et al (1985). Mauchamp et al (2006) identified the production of polycalin a peptide of the midgut that binds to chlorophyllide to form a red fluorescent protein (RFP). The features of polycalin were studied by sequencing, MALDI-TOF and other methods and deduced that it is a 2721 amino acid peptide coded by a 45.5 kb gene.

Material and Method. The present project was studied with the primary intention of identifying, understanding and analyzing the gene expression profiles of the pigment systems of *B. mori* using RNA-Seq technology. To accomplish the above intended objective, next generation transcriptomic sequence data of *B. mori* variant Dazao was extracted from sequence retrieval archives (SRA) of NCBI. This sequence was then subjected to the analysis protocol of RNA-Seq Technology. The details of the tools and the raw materials used in this study are provided below:

Extraction of raw sample sequence data. In the present study we have used NCBI's SRA as our source for raw transcriptomic sequence data. The raw transcriptomic sequence from data downloaded from SRA is labeled as follows:

Dazao (Wild type and pigmented). This raw transcriptomic sequence of this variant is of paired-end (PE) type and isolated from the first three integuments of *B. mori* in its fourth molt, 16 hours after head slippage. The details of the extracted raw sample data are mentioned below in Table 1.

Table 1

Details of raw sample sequence data

Sequencing data	
Data Source	NCBI SRA
Source Organism	*Bombyx mori*
Project ID	PRJNA238971
Sample ID	SRR1177795 (DAZAO)
File Size	3.9G bases (DAZAO)
Sequencing Platform	Illumina HiSeq 2000

Extraction of genome sequence data of B. mori. The whole genome sequence data of *B. mori* has been downloaded from SILKDB database in two different file formats – Fasta and Gene Transfer format (GTF). The details of the extracted genome sequence data of *B. mori* are mentioned below in Table 2.

Table 2

Details of the extracted genome sequence data of *Bombyx mori*

Reference genome	
Data Source	SilkDB (http://silkworm.genomics.org.cn)
Source Organism	*Bombyx mori*
File Size	431,707,935 Base pairs
Filetype used	Fasta and GTF

Preparation of the NGS workbench. The Transcriptomic data analysis pipeline used in this project is based on the resequencing RNA-Seq protocol. The requirements for the construction of this pipeline and the workflow of the analysis protocol are provided in Table 3. All the tools assembled to prepare the NGS workbench were obtained from online open data sources referred from Trapnell et al (2012), account on gene expression analysis.

Table 3

Tools installed to create the NGS Pipeline

Tool name	*Version*
SRA Toolkit	2.4.2-1
NGSQC Toolkit	9V2.3.3
TopHat	2.0.3.linux_X86_64
Bowtie	2-2.2.3-linuxX86_64
Cufflinks Package	1.3.0-1
CummerBund	2.10.0
Blast2GO	3.0.9

Transcriptome resequencing data analysis pipeline. This analysis protocol followed for this study keeping in mind its intended objective is the RNA-Seq protocol for transcriptome resequencing (Wolf 2013). This protocol is one of the highly cited analysis protocols of NGS technologies and the work outline of this protocol is mentioned below in Figure 1.

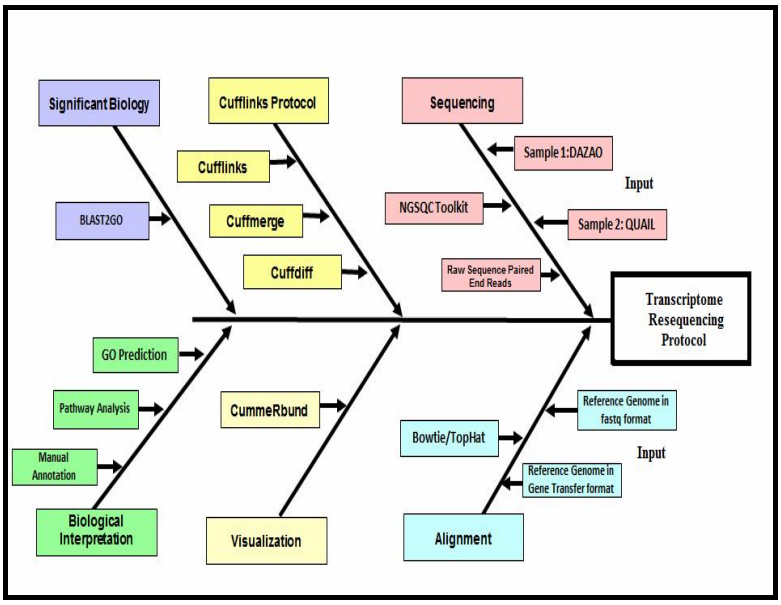


Figure 1. Workflow of transcriptome resequencing protocol.

Results and Discussion. The present study focused on identifying the interplay of genes involved in production of pigments in *B. mori*. In this project, the NGS sequence data of *B. mori* variant – Dazao (wild type and pigmented) was studied for gene level associations and their impacts on pigment production. To accomplish this task, the transcriptome sequence data from the sample was analyzed using the RNA-Seq protocol. A detailed report on the results obtained during the analysis is provided in this section.

Quality check (QC) of raw sample data. The quality of the NGS sample sequence data was tested by the software tool NGSQC Toolkit. NGSQC toolkit uses phred score of the sequence data as a parameter to understand the quality of the raw sample sequence data on the basis of the overall quality, base position, base composition and quality distribution. In this study, the reads of the NGS raw sample sequence data were classified as high quality reads based on a phred score cutoff of 30. The QC results obtained from NGSQC Toolkit for Dazao is as shown in Figure 2.

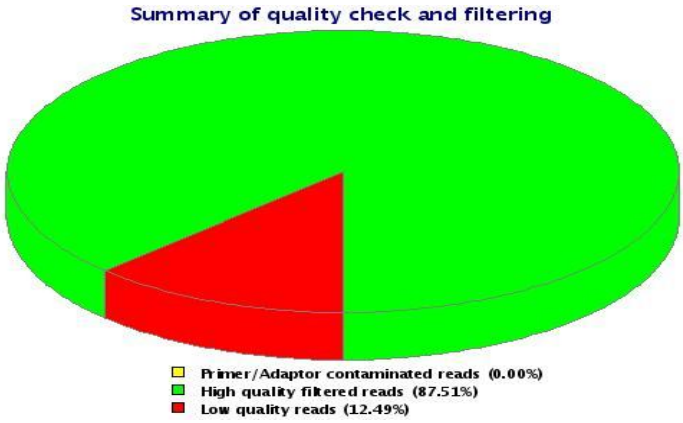


Figure 2. Overall quality of sequence data of *Bombyx mori* variant Dazao.

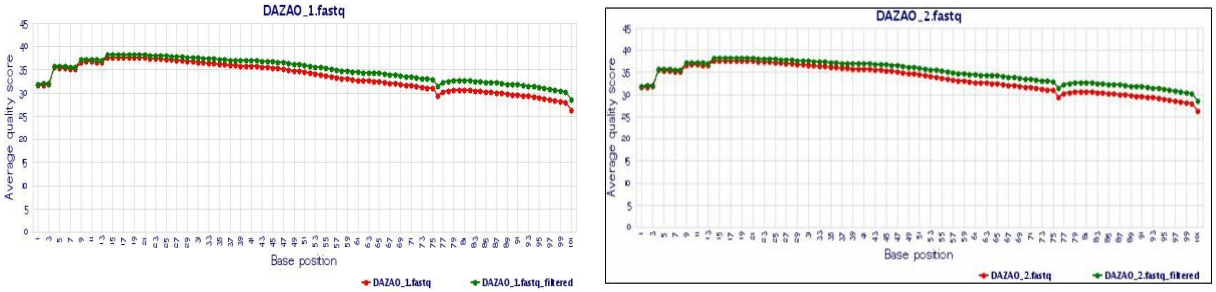
The study of the quality control reports for Dazao indicated that the sample had good sequence data quality based on the following criteria:

a) Overall quality of sequence data – This step provided an overview of the read quality of the entire sample data and classified reads above the phred score 30 as high quality and the remaining as low quality reads. In this study, Dazao passed the QC test wherein 87.51% of the reads were of high quality.

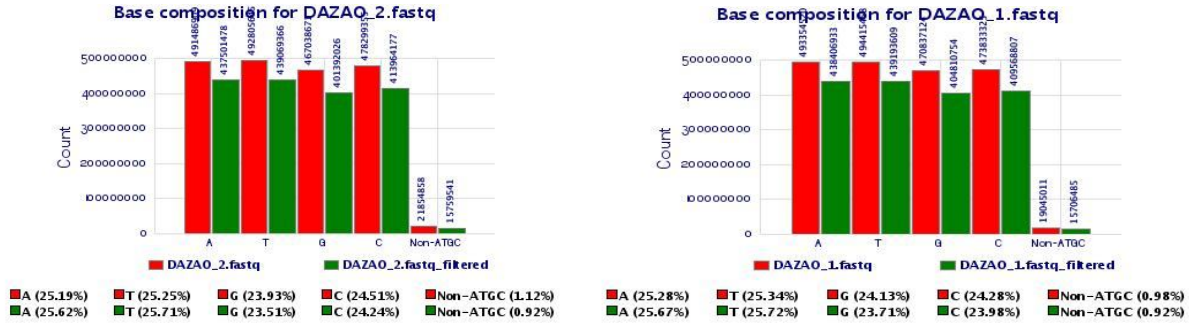
b) Sample sequence quality based on base position – This step analyses the sample quality on the basis of its position at each base, only samples with phred score greater than 30 at each base position pass this QC parameter. Figures 3 of our study for base position indicate most of the bases at each base position have a phred score above 30 in our sample Dazao.

c) Sample sequence quality based on base composition - This step analyses the sample quality on the basis of nucleotide composition of the sequence data. A good quality sample sequence data must have minimal non-ATGC content in it. Figures 3 for base composition indicate the non-ATGC content of our sample to be as minimal as 1% of the nucleotide composition of sample content.

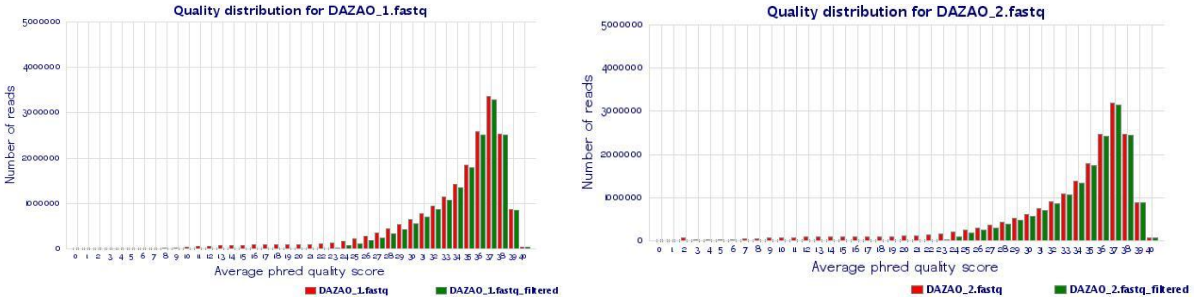
d) Sample sequence quality based on quality distribution – This step analyses the sample quality on the criteria that more than 70% of the sample sequence reads must have a phred score above 30 invariably representing itself as a right kurtosis graph. Figures 3 for quality distribution indicate that more than 70% of both our sample reads have a phred score more than 30.



Assessment of sequence data quality of Dazao based on base position.



Assessment of sequence data quality of Dazao based on base composition.



Assessment of sequence data quality of Dazao based on quality distribution.

Figure 3. Quality control report for *Bombyx mori* variant Dazao.

Reference guided alignment of sample sequence data with genome of B. mori. The sequence file of the sample Dazao was aligned to the reference genome using the Tophat and Bowtie alignment tools. The alignment profiles for the sample are summarized below in Figure 4.

Based on the alignment output obtained for Dazao the following inferences were made:

a) Concordant alignment – This parameter of the read alignment step was used to assess base to base exact alignment. Alignment output indicates that Dazao shows 67.60% concordant alignment.

b) Discordant alignment - This parameter of the read alignment step assess error occurred in exact base to base alignment. Alignment output indicates that Dazao shows 0.4% discordant alignment.

c) Multiple alignment – The parameter of the read alignment step assess the alignment of a single read of exact base to base alignment in more than 20 places of the reference genome. Alignment output indicates that Dazao shows 5.7% multiple alignment.

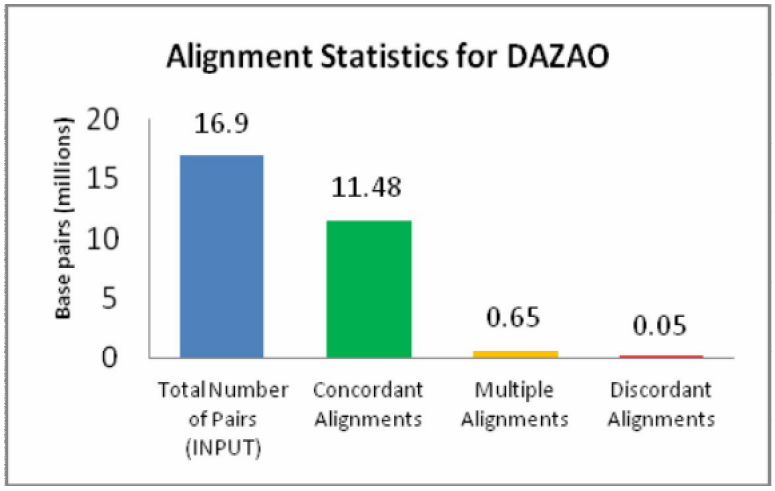


Figure 4. Alignment summary statistics for *Bombyx mori* variant Dazao.

However, the low concordant alignment seen with respect to our sample data can be justified with the fact that the reference genome of *B. mori* is not completely sequenced and still has portions of unassembled scaffolds. Thus the low concordant alignment rate of 67.60% for Dazao sample to the reference genome is considered acceptable.

Cufflinks protocol. The cufflink protocol uses the information from alignment BAM files generated from read alignment step to analyze transcripts for gene or transcript variant discovery. The cufflink protocol comes in the form of package of three tools – cufflinks, cuffmerge and cuffdiff. These three tools are used in transcript compliation, gene identification and gene expression analysis of our sample transcript data.

a) ***Transcript compliation***

In this step Dazao sample transcript reads aligned to *B. mori* genome were complied into assemblies utilizing the Cufflinks tool. The output of this step is obtained in the form of a gtf file. This gtf is sectioned into assemblies providing important information on:

- Chromosome/Scaffold ID
- Classification (exon/transcript)
- Length of the transcript/gene
- FPKM values

b) ***Gene identification***

This step involves merging of unique and common transcript fragments belonging to the sample analyzed during transcriptome analysis into a single gtf file. This step is mainly performed to prevent data redundancy and allow accurate prediction using the cuffmerge tool of cufflinks protocol. In this study, cuffmerge tool was used to merge the unique and

common transcript fragments belonging to Dazao variant of *B. mori* in a single file called merged.gtf. After the cuffmerge step, a total of 1,67009 transcripts were identified from the sample in the gtf file. This merged.gtf file forms the base of all the biological analysis performed henceforth in the RNA-Seq protocol.

c) ***Gene expression***

The step involves the study of gene expression, expression analysis and identification and classification of transcripts into novel, known and isoform transcripts. This step is accomplished on the basis of output files generated from cuffdiff tool. In this study, the non-redundant fields of the cuffdiff output files were extracted into a Microsoft Excel file. The transcript data was then filtered based on the following criteria:

a) Fold expression – It is a measure to check if a particular transcript is expressed or not. All transcripts with FPKM value greater than one were considered to be expressed and taken forward for further analyses.
b) Significant test – It is a test were only those transcripts with p-value greater than 0.05 and q-value greater than 0.1 were considered to show significant expression and taken forward for further analyses.

Based on the above two criteria, a total of 17,253 transcripts were filtered out to be expressed transcripts from our study transcript data.

This step also helped in the classification of transcripts based on class code key:

u: Unknown, intergenic transcript

j: Potentially novel isoform (fragment) - at least one splice junction is shared with a reference transcript

=: Complete match of intron chain

In this study, class code based classification of transcripts identified a total of 5259 transcripts as known for Dazao. A total of 5250 transcripts were identified as potential isoforms for Dazao. And a total of 3401 transcripts were categorized as novel transfrags for Dazao (Figure 5).

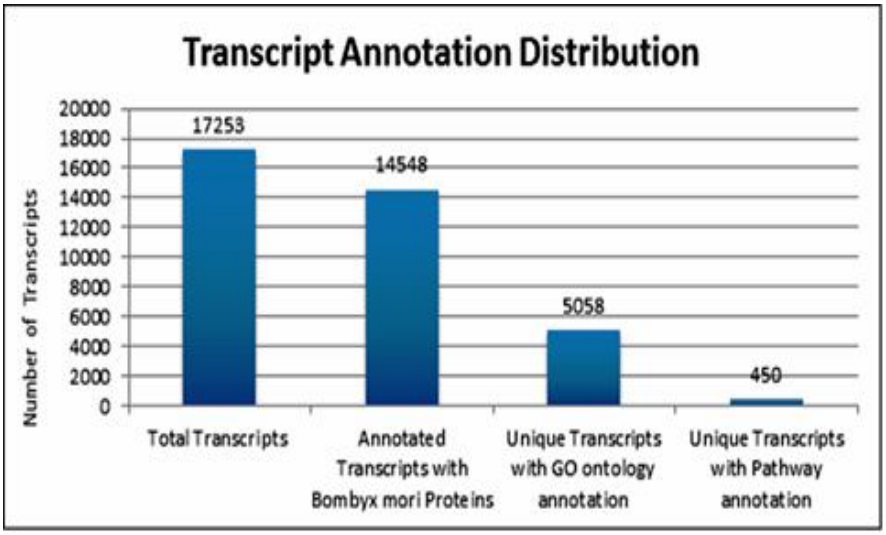


Figure 5. Overall view of the annotation distribution among transcriptome data.

Determination of GO/pathway. In this step, the total number of transcripts identified from Dazao variant of *B. mori* were put together in a file called Tconslist. An attempt was made to identify the functional and pathway annotation of the 17,253 transcripts listed in Tconslist from ENZYME, KEGG and *B. mori* protein annotation derived from GO and NRDB databases with the help of Blast2GO tool. This tool was successful in identifying the functional and pathway annotations for 14,548 transcripts. The Figure 6 below provides, the overview of functional and pathway annotation distribution among the transcripts.

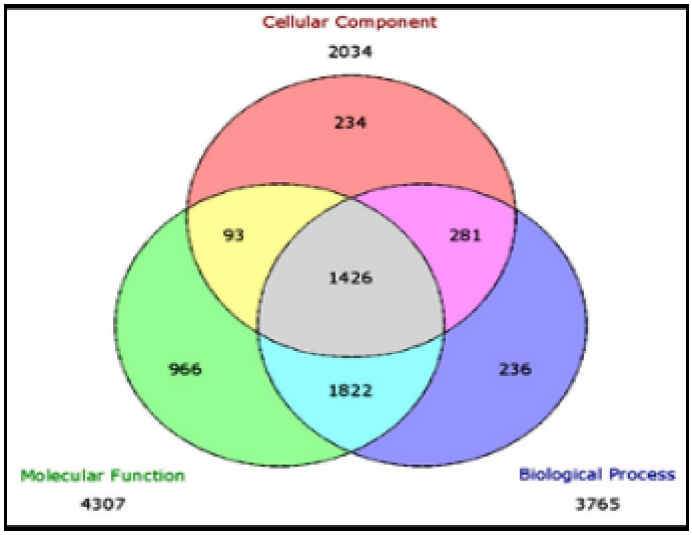


Figure 6. Overview of ontology based distribution of transcriptome data.

The above Figure 6 summarizes that, in this study Blast2GO tool was used to annotate the 17,253 expressed transcripts identified from Cufflink protocol analysis of Dazao transcriptome data. Out of which the tool was successful in identifying the functional and pathway annotations for 14,548 transcripts. The Blast2GO reported 5058 transcripts with unique gene ontology annotation and 450 transcripts with unique pathway annotation. This step also quantifies and distributes the transcripts on the basis of gene ontology terms. The 5058 transcripts annotated with gene ontology terms were classified on the basis of their function, wherein 2034 transcripts were found to be involved in cellular component production, 4307 transcripts were found to be expressing for molecular functions of *B. mori* and a total of 3765 transcripts were set to take part in biological processes. About 1426 transcripts were found to be participating in cellular component, molecular function and biological processes. A detail on the split of functionality of transcripts based on GO terms is as shown in Figure 5.

The functional annotations identified from Blast2GO tool were mapped against KEGG pathway database. This method grouped each transcript into different metabolic pathways based on the proteins or functionality elucidated to each transcript after the BLAST search. The functionally annotated 14,548 transcripts in our study were found to be involved in 99 different metabolic pathways, wherein purine metabolism was highlighted to be the top expressed metabolic pathway with 83 unique participating transcripts as shown in Figure 7.

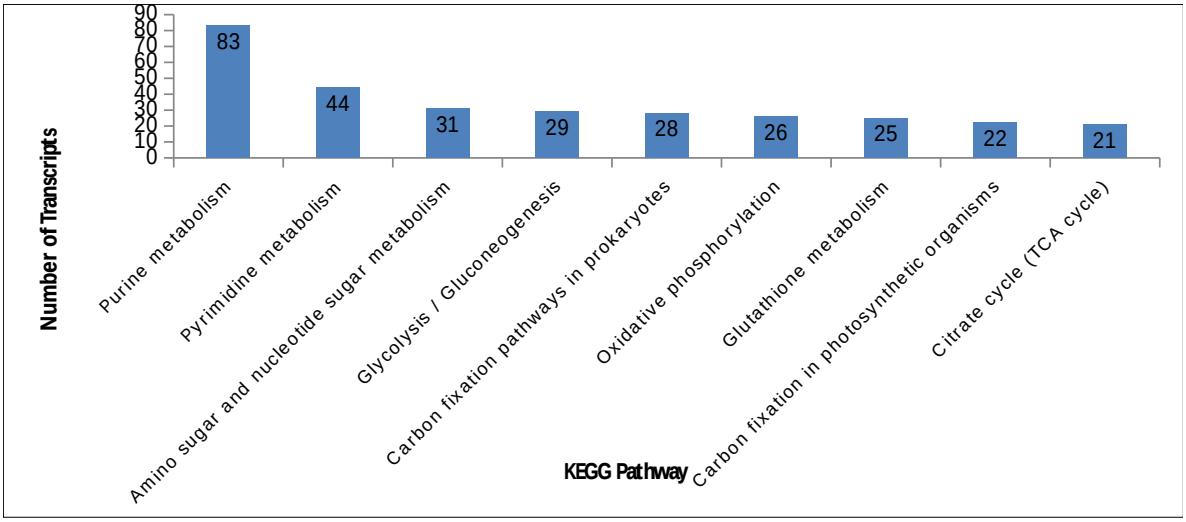


Figure 7. Top expressed pathways based on transcriptome data set.

Biological interpretation. Our study involved exploration of gene and transcript variants associated with pigment production in *B. mori*. Microsoft Excel was used to filter out transcripts involved in pigment production based on functional annotation derived from GO and NRDB databases and metabolic pathway associations which are obtained from a search against KEGG pathway database. These transcripts were further manually annotated with KEGG pathway database to outline pathways along with associated enzymes involved in pigment production. Based on manual and software aided annotation, four pigment pathways were identified and their gene expression was studied in Dazao variant of *B. mori.* The details of which are given below:

a) **Melanins** – these black pigments were found to be produced as a product of tyrosine metabolism. Here the gene **BMG003866** coding for enzyme tyrosine hydrogenase was used to convert phenylalanine to tyrosine was converted to Dopaquinone with the help of the enzyme prophenoloxidase coded by the gene **BMGA013116** which eventually is converted to melanin through spontaneous reactions. The above theory is illustrated in the Figure 8.

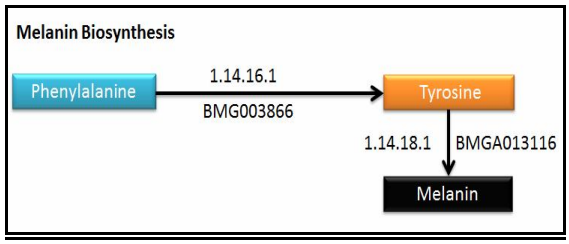


Figure 8. Melanin production pathway.

b) **Ommochromes** - these brown pigments were found to be produced as a product of tryptophan metabolism. Here the gene **BMGA006968** coding for enzyme kynuerine-3-monooxygenase was used to convert tryptophan to 3-hydroxyl-L-kynuerine. This inturn was converted to ommochromes with the help of a pyridoxal protein coded by the novel isoform gene tentatively labeled as **BMGX**. The above theory is illustrated in Figure 9.

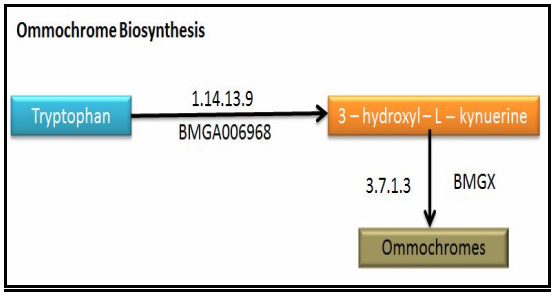


Figure 9. Ommochrome production pathway.

c) **Sepiaterin** – these florescent yellow pigments were found to be produced by the GTP obtained as a product of pyridine metabolism. Here the gene **BMGA001235** coding for enzyme GTP cyclohydrase was used to convert GTP to H2-neoptrin-phosphate. This inturn was converted to 6-Pyruvoy-H4-pterin with the help of the enzyme 6 pyruvoy – H4-pterin synthethese coded by **BMGA003643**. 6-Pyruvoy-H4-pterin was then converted to sepiaterin based on the enzyme sepiaterin reductase expressed by the gene **BMGA009103.** The above theory is illustrated in the Figure 10.

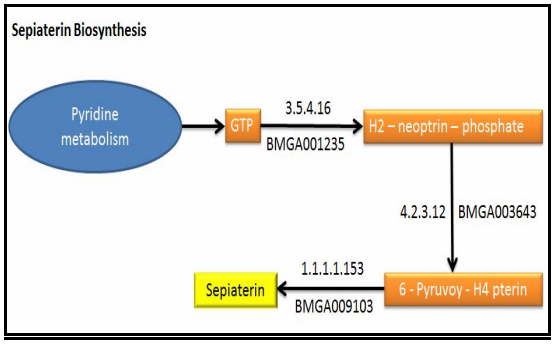


Figure 10. Sepiaterin production pathway.

d) Betalain - these pinkish red pigments were found to be produced as a product of tyrosine metabolism. Here the gene **BMGA013116** coding for enzyme prophenoloxidase was used to convert tyrosine to L-DOPA. This inturn was converted to betalmic acid with the help of the enzyme monophenol monooxygenase coded by **BMGA007234**. Also L-DOPA was converted to dopaquinone and then converted to Amaranthin based on the enzyme synthetase expressed by the gene **BMGA002936.** The above theory is illustrated in the Figure 11.

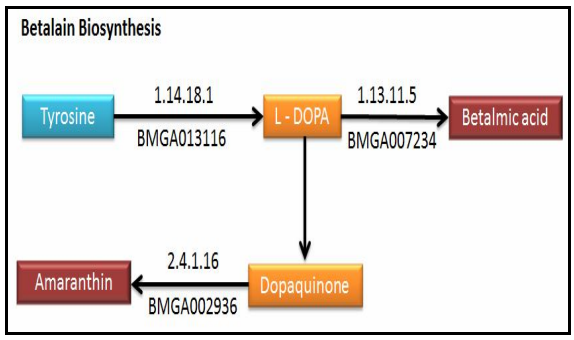


Figure 11. Betalain production pathway.

The nomenclature of enzymes involved in pigment production pathways represented in Figure 8 - 12 are outlined in the Table 4. The expression profiles of genes involved in pigment pathways represented in Figure 8-11 were calculated to understand the likeliness of Dazao variant to express the pigments naturally. Figure 12 shows the variation in the gene expression of Dazao for the production the various pigment pathways hypothesized in Figure 8–11.

Table 4

Legend of enzymes represented in pigment pathways of Figure 8-11

Pathway	*Enzyme code*	*Enzyme nomenclature*
Melanin production	1.14.16.1	Tyrosine hydrogenase
	1.14.18.1	Prophenoloxidase
Ommochrome production	1.14.13.9	Kynuerine 3 monooxygenase
	3.7.1.3	Pyridoxal protein
Sepiaterin production	3.5.4.16	GTP cyclohydrase
	4.2.3.12	6 pyruvoy – H4 pterin synthetase
	1.1.1.153	Sepiaterin reductase
Betalain production	1.14.18.1	Prophenoloxidase
	1.13.11.5	Monophenol monooxygenase
	2.4.1.16	Synthetase

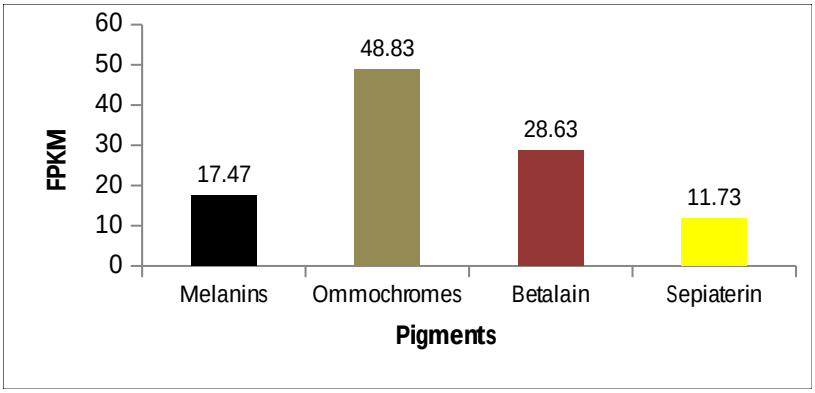


Figure 12. Differential gene expression analysis of pigments produced.

Conclusions. The outcome of this study indicated that Dazao variant of *B. mori* was capable of production of melanins, ommochromes, sepiaterin and betalain. It can be hypothesized that this variant of *B. mori* can be utilized for higher pigment production in conjugation with several chemical components, which might act as enhancers to improve the gene expression of pigment gene networks leaving a huge scope for future work within this research topic. Also the co-relation between pigments on the silkworm body to the color of silk produced can be studied on genomic backgrounds that eventually might lead to better quality and surplus production of naturally colored silk. Lastly, information derived from this study can be assembled to improve the annotation content of the entire Lepidopteran group.

Acknowledgements. Authors would like to acknowledge, Dr. K.N. Subramanya, Principal, Dr. Pushpa Agrawal, HoD, Department of Biotechnology, Dr. Vidya Niranjan, Professor, Department of Biotechnology, RV College of Engineering for their help rendered during the course of work.

References

Adams D. S., Eickbush T. H., Herrera R. J., Lizardi P. M., 1986 A highly reiterated family of transcribed oligo(A)-terminated, interspersed DNA elements in the genome of *Bombyx mori*. J Mol Biol 187(4):465-478.

Busnel R., Levy G., Polonovski M., 1950 Identity of pterine pigment with blue fluorescence of *Bombyx mori* and fluorescyanine. C R Seances Soc Biol Fil 144(5-6):334-335.

Chen J., Wang Q., Hua Z. Z., Du G. C., 2007 Research and application of biotechnology in textile industries in China. Enzyme Microb Technol 40:1651-1655.

Drilhon A., 1951 Chromatographic and chemical identification of a fluorescent substance in the blood and in the malpighian tubules of the larvae of *Bombyx mori* affected with polyhedral disease. C R Hebd Seances Acad Sci 232(20):1876-1878.

Drilhon A., Vago C., 1953 Modifications in the chromatographic form of free amino acids and fluorescent substances in the hemolymph of *Bombyx mori* L. following a paralysis of bacterial origin. Experientia 9(4):143-145.

Duan J., Li R., Cheng D., Fan W., Zha X., Cheng T., Wu Y., Wang J., Mita K., Xiang Z., Xia Q., 2010 SilkDB v2.0: A platform for silkworm (*Bombyx mori*) genome biology. Nucleic Acids Res 38(Databaseissue):D453-456. doi: 10.1093/nar/gkp801.Epub 2009 Sep 30.

Gangopadhyay D., 2008 Sericulture industry in India – A review. In: India, science and technology. Report. Banerjee P. (ed), pp. 334-339, S&T for Rural India and Inclusive Growth.

Hirata Y., Nakanishi K., Kikkawa H., 1950 Xanthopterin obtained from the skins of the yellow mutant of *Bombyx mori* (silkworm). Science 111(2892):608-609.

Hirata Y., Nawa S., 1951 Fluorescyanine (ichthyopterin) obtained from eggs of *Bombyx mori* and carp scales. C R Seances Soc Biol Fil 145(9-10):661-663.
Hirata Y., Nawa S., Matsuura S., Kakizawa H., 1952 Synthesis of pterins and a note on the constitution of a pterin of *Bombyx mori*. Experientia 8(9):339-341.
Iizuka T., Sezutsu H., Tatematsu K.-I., Kobayashi I., Yonemura N., Uchino K., Nakajima K., Kojima K., Takabayashi C., Machii H., Yamada K., Kurihara H., Asakura T., Nakazawa Y., Miyawaki A., Karasawa S., Kobayashi H., Yamaguchi J., Kuwabara N., Nakamura T., Yoshii K., Tamura T., 2013 Colored fluorescent silk made by transgenic silkworms. Adv Funct Mater 32(3):7-10.
Ishiguro I., Nagamura Y., 1971 Red pigments in the urine of the normal silkworm and rb-mutant moths. Igaku To Seibutsugaku 82(4):149-153.
Ishiguro I., Nagamura Y., Yamamoto T., 1971 Formation of phenoxazine pigment and conversion of 3-oxy-kynurenine during metamorphosis of silkworms. Seikagaku 43(7):401-407.
Jimbo G., Nishiyama T., 1962 Embryonic studies on the fluorescent compounds in silkworms. Bull Yamaguchi Med Sch 9:81-84.
Jin-Shan X., Qing-You X., Jun L., Guo-Qing P., Ze-Yang Z., 2005 Survey of long terminal repeat retrotransposons of domesticated silkworm (*Bombyx mori*). Insect Biochem Mol Biol 35(8):921-929.
Kuan K. N., Lee Y. Y., Melius P., 1979 An improved method of light-induced pigmentation. Biochem J 177(3):981-983.
Li Y., Wang G., Tian J., Liu H., Yang H., Yi Y., Wang J., Shi X., Jiang F., Yao B., Zhang Z., 2012 Transcriptome analysis of the silkworm (*Bombyx mori*) by high-throughput RNA sequencing. PLoS One 7(8):e43713. doi: 10.1371/journal.pone.0043713.
Matsumoto S., Isogai A., Suzuki A., 1985 N-terminal amino acid sequence of an insect neurohormone, melanization and reddish coloration hormone (MRCH): heterogeneity and sequence homology with human insulin-like growth factor II. FEBS Lett 189(1):115-118.
Mauchamp B., Royer C., Garel A., Jalabert A., Da Rocha M., Grenier A. M., Labas V., Vinh J., Mita K., Kadono K., Chavancy G., 2006 Polycalin (chlorophyllid A binding protein): a novel, very large fluorescent lipocalin from the midgut of the domestic silkworm *Bombyx mori* L. Insect Biochem Mol Biol 36(8):623-633.
Ogura T., Okano K., Tsuchida K., Miyajima N., Tanaka H., Takada N., Izumi S., Tomino S., Maekawa H., 1994 A defective non-LTR retrotransposon is dispersed throughout the genome of the silkworm, *Bombyx mori*. Chromosoma 103(5):311-323.
Quan G. X., Kim I., Kômoto N., Sezutsu H., Ote M., Shimada T., Kanda T., Mita K., Kobayashi M., Tamura T., 2002 Characterization of the kynurenine 3-monooxygenase gene corresponding to the white egg 1 mutant in the silkworm *Bombyx mori*. Mol Genet Genomics 267(1):1-9.
Randolph V. Lewis, 2006 Spider silk: ancient ideas for new biomaterials. Chem Rev 106(9):3762-3774.
Sakurai S., 1968 Genetical and biochemical studies of chromogranules in larval skin cells of the silkworm. Yokohama Med Bull 19(2):29-43.
Shimomura M., Minami H., Suetsugu Y., Ohyanagi H., Satoh C., Antonio B., Nagamura Y., Kadono-Okuda K., Kajiwara H., Sezutsu H., Nagaraju J., Goldsmith M. R., Xia Q., Yamamoto K., Mita K., 2009 KAIKObase: An integrated silkworm genome database and data mining tool. BMC Genomics 10:486. doi: 10.1186/1471-2164-10-486.
Shvetsova O. I., Zurabova E. R., 1966 Pigment producing variants of *Bacillus thuringiensis* Berliner isolated from the silk worm. Mikrobiologiia 35(4):726-731.
Suzuki M. G., Daimon T., Shinoda T., Taniai K., Asaoka K., Niwa R., Kawaoka S., Katsuma S., Tamura T., Noda H., Kasahara M., Sugano S., Suzuki Y., Fujiwara H., Kataoka H., Arunkumar K. P., Tomar A., Nagaraju J., Goldsmith M. R., Feng Q., Xia Q., Yamamoto K., Shimada T., Mita K., 2013 Large scale full-length cDNA sequencing reveals a unique genomic landscape in a lepidopteran model insect, *Bombyx mori*. G3 (Bethesda) 3(9):1481-1492.

Trapnell C., Roberts A., Goff L., Pertea G., Kim D., Kelley D. R., Piment H., Salzberg S. L., Rinn J.L., Pachter L., 2012 Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7:562-578.

Wolf J. B. W., 2013 Principles of transcriptome analysis and gene expression quantification: an RNA-Seq tutorial. Mol Ecol Resour 13:559-572.

Yamamoto K., Nohata J., Kadono-Okuda K., Narukawa J., Sasanuma M., Sasanuma S., Minami H., Shimomura M., Suetsugu Y., Banno Y., Osoegawa K., de Jong P. J., Goldsmith M. R., Mita K., 2008 A BAC-based integrated linkage map of the silkworm *Bombyx mori*. Genome Biol 9(1):R21. doi: 10.1186/gb-2008-9-1-r21.

Zhao Q., Zhu Z., Kasahara M., Morishita S., Zhang Z., 2013 Segmental duplications in the silkworm genome. BMC Genomics 14:521. doi:10.1186/1471-2164-14-521.

*** International Silkworm Genome Consortium 2008 The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. Insect Biochem Mol Biol 38(12):1036-1045.

Received: 01 April 2016. Accepted: 19 September 2016. Published online: 26 September 2016.
Authors:
Archana Vidhyadhar Pawar, RV College of Engineering, Department of Biotechnology, India, Karnataka, Bengaluru 560059, e-mail: archanapwr23@gmail.com
Aall Hanumanthareddy Manjunatha Reddy, RV College of Engineering, Department of Biotechnology, India, Karnataka, Bengaluru 560059, e-mail: ahmanjunatha@rvce.edu.in

How to cite this article:
Pawar A. V., Manjunatha Reddy A. H., 2016 Transcriptome profiling of genes that govern the pigment systems of *Bombyx mori* variant Dazao. ABAH Bioflux 8(2):39-52.