A Walk Through the Maze of Secondary Metabolism in Orchids: A Transcriptomic Approach

Devina Ghai1, Arshpreet Kaur1, Parvinderdeep S. Kahlon2, Sandip V. Pawar3 and Jaspreet K. Sembi1*

1Department of Botany, Panjab University, Chandigarh, India, 2Chair of Phytopathology, TUM School of Life Sciences, Technical University of Munich, Freising, Germany, 3University Institute of Pharmaceutical Sciences, Panjab University, Chandigarh, India

Orchids have a huge reservoir of secondary metabolites making these plants of immense therapeutic importance. Their potential as curatives has been realized since times immemorial and are extensively studied for their medicinal properties. Secondary metabolism is under stringent genetic control in plants and several molecular factors are involved in regulating the production of the metabolites. However, due to the complex molecular networks, a complete understanding of the specific molecular cues is lacking. High-throughput omics technologies have the potential to fill up this lacuna. The present study deals with comparative analysis of high-throughput transcript data involving gene identification, functional annotation, and differential expression in more than 30 orchid transcriptome data sets, with a focus to elucidate the role of various factors in alkaloid and flavonoid biosynthesis. Comprehensive analysis of the mevalonate (MVA) pathway, methyl-d-erythritol 4-phosphate (MEP) pathway, and phenylpropanoid pathway provide specific insights to the potential gene targets for drug discovery. It is envisaged that a positive stimulation of these pathways through regulation of pivotal genes and alteration of specific gene expression, could facilitate the production of secondary metabolites and enable efficient tapping of the therapeutic potential of orchids. This further would lay the foundation for developing strategies for genetic and epigenetic improvement of these plants for development of therapeutic products.

Keywords: secondary metabolism, transcriptome, orchids, alkaloids, flavonoids

INTRODUCTION

Orchids are members of one of the most advanced plant families, the Orchidaceae with their unique morphology (labellum, gynostemium), functional characteristics, and ecological adaptations (mycorrhizal association, and velamen) that are not found in model plants. Though popular as affluent ornamentals, orchids were first discovered for their therapeutic properties. The restorative properties of orchids have been well documented since times immemorial, Theophrastus in his book named “Enquiry into Plants” reported the use of orchids as therapeutics. These plants have also found reference in Indian and Chinese traditional pharmacopeia. In Indian Ayurvedic system of medicine, “Ashtavarga” is an important formulation, consisting of eight herbs, out of which four are orchids, that is, Habenaria edgeworthii (vriddhi), Habenaria
intermedia (riddhi), Malaxis acuminata (jeevaka), and Malaxis muscifera (rishibhak). Similarly, in Chinese medicine, Anoectochilus roxburghii has been promoted as “King medicine” to treat snake bites, lung and liver disease, and hypertension (He et al., 2006). “Shi-Hu,” an orchid-based therapeutic formulation, prepared from Dendrobium nobile and allied species, is prized as a tonic because of its efficiency in treating lung, kidney, and stomach diseases, hyperglycemia, and diabetes (Bulpitt et al., 2007). “Tian-Ma” derived from tubers of Gastrodia elata is effectively used in the treatment of headaches, migraines, epilepsy, high blood pressure, rheumatism, fever, and nervous problems (Kong et al., 2003). In addition to their use as therapeutics, these plants have also been widely used as tonics and restoratives. The most important example is Dactylorhiza hatagirea which is used as an aphrodisiac (Lawler, 1984). Several orchids, such as Shwethuli (Zeuxine strateumatica) and Salabmisri (Eulophia dabia), Vanda testacea, and Rhynchostylis retusa, are used as aphrodisiacs, blood purifiers, general restorative tonics, and for treating rheumatism, piles, bronchitis, and inflammations (Chauhan, 1990; Vij et al., 2013; Hossain et al., 2020; Figure 1). These healing and restorative properties are due to the presence of a rich diversity of phytochemicals which are bioactive and are responsible for the pharmacognostic potential of these plants (Teoh, 2016).

![Figure 1](image_url)

**FIGURE 1** | Some therapeutically important orchid species. (A), Eulophia dabia (D.Don) Hoeh.; (B), Zeuxine strateumatica (L.) Schltr.; (C), Dactylorhiza hatagirea (D.Don) Soó; (D), Malaxis muscifera (Lindl.) Kuntze; (E), Habenaria intermedia D.Don; (F), Habenariapectinata D.Don; (G), Vanda testacea (Lindl.) Pichř.; (H), Platanthera edgeworthii (Hook.f. ex Collett) R.K.Gupta; (I), Crepidium acuminatum (D.Don) Szlach.

The integration of traditional knowledge with modern research can pave a way as promising leads for the discovery of novel drugs with greater therapeutic potential than synthetic medicine offering new horizons in the field of therapeutics and drug discovery. However, the studies in this direction are not commensurate with the immense potential of these plants. This is mainly due to lack of complete understanding of the spectrum of molecular networks of secondary metabolism. Even though there have been a number of studies on the phytochemical profiling and the biological activity, there is limited information about the regulating molecular cues and the alternate biosynthetic routes which are utilized in these plants as a survival strategy in harsh and dynamic climatic conditions. Various omics approaches can be instrumental to understand and elucidate these complex mazes and help in utilization of these plants as therapeutics to their fullest potential.

Recent times have revolutionized the process of deciphering the genetic identity of the germplasm by using minimal amount of tissues to generate humongous volume of data using transcriptomic approach. Genome editing with the help of transcriptomic sequencing provide extra choices for genetic improvement in orchids. For techniques like CRISPR/Cas9, the sequence of the genome of the host can ascertain the specific and accurate target sites to increase the efficiency of the genome editing process (Kui et al., 2017) and can be highly beneficial for overall improvement of the germplasm. Transcriptomic sequencing has also helped in increasing the pace for the development of Simple Sequence Repeats (SSR), which are the microsatellite markers with random tandem repeats of 2–6 nucleotides. These markers are widely used because of their reproducibility, co-dominant nature, extreme polymorphism, simplicity, abundance, and easy amplification. The development of SSR markers in these medicinally important orchids can help in germplasm breeding, marker-assisted selection, parentage analysis, and genetic diversity studies. The SSR markers identified can help in evaluating and understanding genetic relationships quantitatively and qualitatively (Li et al., 2014) and help in constructing genetic maps of these plants which will further help in taxonomy, genetics, and genomic studies.

The reference genome of many medicinal non-model plants is not available. Transcriptomic approach provides an alternative way for collecting high-throughput data for gene identification, expression analysis, and putative functional characterization using metabolic profiling data (Gongora-Castillo and Buell, 2013). Whole transcriptome shotgun sequencing (WTSS) makes it possible to probe the genes of various metabolite biosynthesis processes and the relationship between the genes and plant metabolites. Another approach, termed as the Phytochemical genomics approach, involves sequence data sets combined with metabolomic data sets to elucidate the complete profile of secondary metabolites. In Digital gene expression analysis, differential expression of genes which are involved in secondary metabolism is studied to decipher the genetic variability and help in the drug discovery. The development of single-cell transcriptomics will aid in identifying networks and pathways and further facilitate drug discovery and development. The
The present study is an exhaustive review of the omics research on secondary metabolism in orchids, primarily focusing on the use of transcriptomic data for the analysis of genes and pathways associated with the synthesis of secondary metabolites and could be further be used for establishing the therapeutic potential of the orchids.

**ESTABLISHMENT OF ORCHIDS AS THERAPEUTIC AGENTS**

The therapeutic potential of orchids has been reported since times immemorial. In 1579, Langham (1579) reported the antipyretic and anti-diarrheal properties of orchids. A Caribbean folklore mentions the use of Vanilla clavicularata for treating wounds and syphilis (Griffith, 1847) while the flowers of Vanilla griffithii (Burkill, 1935) and leaf paste of Vanda roxburghii were used in treating fever (Chawla et al., 1992). Dendrobium huoshanense stems are reported to be beneficial for the eye, stomach, and liver ailments (Hsieh et al., 2008; Luo et al., 2008) while those of Dendrobium moniliforme are reported to be antipyretic (Zhao et al., 2003). Oil-based extracts of stems and leaves of Anoectochilus formosus are effective for the treatment of hypertension, impotency, liver spleen disorders, and chest and abdominal pains (Satish et al., 2003). Leaf decoction of Dendrobium candidum is used for treating diabetes (Wu et al., 2004). Traditional usage of orchids as restoratives and tonics have been widely and commonly reported. The tubers of Dactylorhiza hatagirea have been used for the preparation of "Salep" which possess healing qualities (Lawler, 1984). Similar preparations like “dbang lag” have been used to provide sustenance for Tibetan monks practicing in remote caves (Teoh, 1991). Such studies coupled with ethnobotanical knowledge formed basis of many systematic reviews on utilization of orchids as therapeutic agents (Lawler, 1984).

Due to the significant role of orchids in the traditional medicine system, it has become imperative that these traditional remedies should be utilized for the discovery of new therapeutics. A plethora of studies has been reported ever since, to investigate the role of orchids as promising source of bioactive agents. A number of reports on the antioxidant and anti-inflammatory potential of various orchids like Phalaenopsis hybrids (Minh et al., 2016) and Dendrobium officinale (Zhang et al., 2017) have come up. Cytotoxic and apoptotic effects have also been reported in Dendrobium crepidatum and D. chrysanthum (Prasad and Koch, 2016). Antimicrobial activity has also been documented in Dendrobium moniliforme (Paudel et al., 2018). Antihyperglycemic (Dactylorhiza hatagirea; Choukarrya et al., 2019), anti-diabetic and hepatoprotective activity (Galanthe fimбриата; Peng et al., 2019) have been reported.

To provide a sound scientific scaffolding for development of potential therapeutic products, efforts have been also directed to isolate and profile the phytochemicals from plant extracts. Various classes of secondary metabolites have been isolated from different plant parts and evaluated for biological activity. Phenanthrenes, like denbinobin, from Dendrobium nobile, showed potential cytotoxic activity (Lee et al., 1995), prevented metastatic gastric cancer, and showed potent therapeutic activity against hepatic fibrosis (Yang et al., 2007; Song et al., 2012). Similarly, kinsenoside from Anoectochilus roxburghii showed antihyperglycemic activity (Zhang et al., 2007). Cymbidium A from Cymbidium goeringii is responsible for the hypotensive and diuretic activity (Watanabe et al., 2007). Flavones C-glycosides and anthocyanins from red Phalaenopsis hybrids exhibited antioxidant activity (Kuo et al., 2010). Polysaccharides from Dendrobium officinale (Liu et al., 2011) and Gastrodia elata (Bao et al., 2017) have exhibited immune-enhancing potential. Galactoxyloglucan (GXG), a purified polysaccharide from Dendrobium huoshanense, improved insulin sensitivity, thus preventing hyperglycemia (Wang et al., 2019). Role of flavonoids especially rutin, in imparting antioxidant potential have also been highlighted in Dendrobium officinale (Zhang et al., 2017). Flavonoids of Dactylorhiza hatagirea also exhibited antihyperglycemic activity (Choukarrya et al., 2019). Sesquiterpenoids from Dendrobium nobile exhibited neuroprotective activity (Ma et al., 2019b), while bibenzyl compounds from Dendrobium officinale showed cytotoxic activity (Ren et al., 2020). A group of compounds (phenanthrenes, bibenzyls, glucosyloxybenzyl succinate derivatives, flavonoids, lignans, terpenoids, etc.) isolated from Pleione, showed antitumor, anti-neurodegenerative, and anti-inflammatory biological activities (Wu et al., 2019). Despite a large number of reports on the phytochemical profiling in orchids, the studies are not commensurate with the immense potential of orchids as therapeutic agents. Omics techniques offer a great opportunity to provide an alternate and efficient method to study and characterize specific phytochemicals. Transcriptomic approaches can generate insights to the secondary metabolite biosynthetic pathways and can aid in functional characterization of their key regulatory genes.

**TRANSCRIPTOMIC DATASETS IN ORCHIDS**

Undeterred by the peculiarity in their unique characteristics, orchids are depreciated with respect to understanding their molecular complexities. A complete understanding of the spectrum of the molecular networks by isolated analyses of gene families is not plausible due to the limited availability of orchid genomes. On the other hand, transcriptome-wide analyses can help resolve complex metabolic pathways which are at play in these plants. Transcriptome is a complete set of mRNA and non-coding RNA produced by a cell or organism at a particular point of time. It generates large-scale transcripts that could help in analyzing different gene families all at once and could also guide toward understanding cross-links in mechanisms involved. The analysis begins with the collection of the desired tissue and subsequent isolation of RNA from the collected sample. The isolated RNA is used for the synthesis of complementary DNA which is eventually utilized for the construction of libraries after sequencing. There are large numbers of sequencing techniques that are prevalent nowadays, such as Roche/454, Illumina, Applied Biosystems SOLiD, and
Helicos HeliScope (Magi et al., 2010). Even though these techniques produce abundant short reads at a much higher throughput than any Sanger sequencer but data presented after such analysis is a set of short reads composed of several hundred base pairs. The reads, thus, obtained are curated as raw reads. These read are first filtered and adjusted based on the quality control measures. Then the filtered reads are first either reconstructed using de novo assembly in absence of reference genome or assembled by alignment to the reference genome (Wolf, 2013). The assembly of the reads can be performed with tools like Trinity (Grabherr et al., 2011), Velvet (Zerbino and Birney, 2008), SPAdes (Bankevich et al., 2012), or SOAPdenovo-Trans (Xie et al., 2014). The assembled reads form contigs or singletons; both of these are part of unigenes. The functional annotation of the unigenes or transcripts is completed using various databases, such as NCBI,\(^1\) KEGG,\(^2\) and SwissProt.\(^3\) Additionally, the number of reads for a transcript provides the level of its abundance, thus serving as the starting point for biological inference of spatiotemporal gene expression (Wolf, 2013; Ma et al., 2019a). Transcriptome helps in identification of transcripts involved in primary and secondary metabolism and their splice variants (Wang et al., 2009). Comparing the levels of differentially expressed genes at different developmental stages or environmental conditions, provide insights into the physiological status of the tissue at a specific time. These data sets also contain information of small RNAs, long non-coding RNAs, and molecular repeats etc., and provide a tentative framework for functional assertion for putative annotations. These data can serve as an important lead for modern pharmaceutical industry toward development of herbal-based medicines.

High-throughput transcriptomic approaches produce extensive data sets that can be applied to identify candidate key genes in specific physiological processes using co-expression networks analysis (Carrera et al., 2009; Windram et al., 2014). On the other hand, targeted sequencing using degenerate primers proves to be economical and enables exhaustive analysis of specific genes. Specific genes exhibiting significant sequence similarity with genes involved in similar biological processes can be amplified by degenerate primers in related organisms (Wei et al., 2003). Functional validation of putative genes using metabolic profiling of flavonoids using gene-insertion mutants and transgenic plants with overexpressing genes could be used to understand the role genes in secondary metabolism. Further, recombinant proteins and in vitro biochemical assays could be used to decipher catalytic activity of the proteins. This “reverse genetics” approach for gene identification is very promising where bioinformatic prediction of candidate genes preceded the experimental analysis.

There have been a limited number of transcriptome-wide studies in orchids to explore and elucidate different aspects of orchid development (Table 1), however, the efforts are not in line with the immense advantage of using transcriptomic techniques to decipher various molecular networks. The therapeutic potential of orchids is closely associated with the intricate maze of secondary metabolism pathways and their by-products is mainly responsible for their diverse therapeutic properties. These pathways are, in turn, under strict control of an array of molecular factors which regulate the synthesis of phytochemicals. A large number of gene families are specifically associated with various biosynthetic pathways. Transcriptomic data emerging from various studies conducted in orchids have been tabulated in Table 1 and it is evident that Illumina sequencing was the most commonly used sequencing method and Trinity was the most common assembler software used. A maximum number of final reads were obtained in Dendrobium officinale (81,284,898; Yuan et al., 2020) and highest number of unigenes were identified in Dendrobium huoshanense (499,190, Zhou et al., 2020). A huge variation was noticed in the total number of unigenes as reported in different plant parts using different techniques. In Dendrobium officinale, the range in the number of unigenes was observed from 2,99,107 (Shen et al., 2017) to 23,131 (Adejobi et al., 2021) as reported from various tissues. Similarly, in Dendrobium catenatum, 23,139 unigenes were reported from stem tissue (Lei et al., 2018) and the number drastically increased to 478,361 in Dendrobium huoshanense when roots and leaves were also included for analysis (Yuan et al., 2018). This can be attributed to specific gene expression in tissues at various stages of growth and development and environmental conditions. In Phalaenopsis amabilis, a comparative number of unigenes were reported, 37,723 and 34,020, from petals and labellum, respectively (Yang et al., 2014), indicating that a similar genetic profile can be seen in tissues at comparable physiological stages. In Anoectochilus roxburghii, 186,865 unigenes were reported from root, stem, and leaves (Chen et al., 2020). Interestingly, different techniques and platforms used for sequencing analysis can also play a role in this variation. Root, stem, and leaf tissues of Dendrobium huoshanense reported 4,99,190 unigenes when the Illumina HiSeq2000 platform was used (Zhou et al., 2020) while 4,78,361 unigenes were identified when Illumina Hiseq 2500 platform was used (Yuan et al., 2018). Hence, it can be concluded that a lot of variation is observed in the transcriptomic data, and hence, the analysis needs to be supported with substantial functional studies.

**FUNCTIONAL ANNOTATION OF SECONDARY METABOLISM SPECIFIC GENES**

Transcriptomic data can provide a basic lead for functional studies if a unified, systematic, and statistically significant approach is adopted for its assembly and characterization. To scrutinize the functionality of the unigenes identified from the transcriptomic data set, their assessment was carried out against different databases like KEGG, Swissprot, and non-redundant database (Nr; Table 2). The highest similarity of the unigenes was found against the Nr database except in

\(^1\)https://www.ncbi.nlm.nih.gov/
\(^2\)https://www.genome.jp/kegg/
\(^3\)https://www.expasy.org/resources/uniprotkb-swiss-prot
| Plant name | Sequencing platform | Assembly | Plant part | Raw reads | Final reads | Total unigenes | References |
|------------|---------------------|----------|------------|-----------|-------------|----------------|------------|
| Anoectochilus roxburghii | Illumina HiSeq X Ten | Trinity | Roots, Stems, and Leaves | 61,226,728 | 60,425,910 | 67,559,786 | Chen et al., 2020 |
| | Illumina HiSeq 2000 | Trinity v.2.0.6 software | Non-mycorrhizal plant (NM) | 65,007,376 | 64,859,884 | 66,965,132 | Zhang et al., 2020a |
| Bletilla striata | Illumina HiSeq 4000 | Trinity | Leaves, tubers and roots | 270,734,628 | – | – | Ma et al., 2021 |
| Bletilla striata (Thunb.) Reichb. f. varieties | Illumina HiSeq 2000 platform | Trinity | Pseudobulbs | 55,632,192 | 55,492,010 | 65,007,376 | 64,859,884 | 66,965,132 | Chen et al., 2021 |
| Calanthe tsoongiana | Illumina HiSeq X Ten | Trinity | Four transitional stages from seed to seedling | 592,645,857 | 577,527,375 | 73,528 | Jiang et al., 2021a |
| Cymbidium goeringii | Illumina HiSeq™ 2000 platform | Trinity | Floral bud, Half-flowering, Full flowering stage | 161,763,530 | 159,616,374 | 85,868 | Ramya et al., 2019 |
| Cymbidium kanran | Illumina HiSeq™ 2,500 | Trinity | Buds and flowers | 359,645,857 | 577,527,375 | 73,528 | Jiang et al., 2021a |
| Cymbidium longibracteatum | Illumina HiSeq 2000 platform | Trinity | Yellow leaves (YL) | 39,557,830 | 38,356,724 | 5,536,274 | 5,503,245,825 | Zhou et al., 2018 |
| | | | Green leaves (GL) | 39,557,830 | 38,356,724 | 5,536,274 | 5,503,245,825 | Zhou et al., 2018 |
| Cymbidium tortisepalum var. longibracteatum cultivars | Illumina HiSeq 2000 platform | Trinity | Green Rhizome (GR) | 39,557,830 | 38,356,724 | 5,536,274 | 5,503,245,825 | Zhou et al., 2018 |
| | | | Yellow Rhizome (YR) | 39,557,830 | 38,356,724 | 5,536,274 | 5,503,245,825 | Zhou et al., 2018 |
| Dactylorhiza hatagirea | Illumina GA IIx platform | SOAP denovo-Trans | Leaves (L) | 22,009,740 | 21,263,988 | 15,917,274 | 15,456,424 | 26,868,355 | Dhiman et al., 2019 |
| | | | Shoots (S) | 22,009,740 | 21,263,988 | 15,917,274 | 15,456,424 | 26,868,355 | Dhiman et al., 2019 |
| | | | Tubers (T) | 22,009,740 | 21,263,988 | 15,917,274 | 15,456,424 | 26,868,355 | Dhiman et al., 2019 |
| | | | Stems | 22,009,740 | 21,263,988 | 15,917,274 | 15,456,424 | 26,868,355 | Dhiman et al., 2019 |
| Dendrobium catenatum | Illumina HiSeq™ 4000 platform | Trinity | Roots, Stems, and Leaves | 476,746,678 | 444,999,698 | 499,190 | Zhou et al., 2020 |
| Dendrobium huoshanense | Illumina HiSeq 2000 platform | Trinity | Roots, Stems, and Leaves | 736,904,076 | 716,634,006 | 478,361 | Yuan et al., 2018 |
| Dendrobium Nestor (Dendrobium parishii × D. anosmum) | Illumina HiSeq™ 4000 platform | Trinity | Flower bud stage (F) | 50,047,108 | 48,759,280 | 48,759,280 | 47,538,849 | 161,228 | Cui et al., 2021 |
| | | | Half bloom stage (H) | 50,047,108 | 48,759,280 | 48,759,280 | 47,538,849 | 161,228 | Cui et al., 2021 |
| | | | Full bloom stage (B) | 50,047,108 | 48,759,280 | 48,759,280 | 47,538,849 | 161,228 | Cui et al., 2021 |
| | | | Stems | 50,047,108 | 48,759,280 | 48,759,280 | 47,538,849 | 161,228 | Cui et al., 2021 |
| Dendrobium nobile | Illumina HiSeq 4000 platform | Trinity | – | – | – | – | – |

(Continued)
### TABLE 1 | Continued

| Plant name | Sequencing platform | Assembly | Plant part | Raw reads | Final reads | Total unigenes | References |
|------------|---------------------|----------|------------|-----------|-------------|----------------|------------|
| Dendrobium officinale | HiSeq™ 2500 Illumina | – | Roots Control (CK) | 83, 206, 690 | CK | 81,284,898 | 23,131 | Adejobi et al., 2021 |
| | | MeJa treated (MeJa) | 82,623,796 | MeJa | 81,047,188 | | | |
| | | Leaves | – | – | – | | | |
| | BGISEQ-500 Trinity | | Protocorm like bodies and Leaves | 771,499,974 | 747,574,430 | 24,927 | | |
| | Illumina 4000 platform | | Roots, Stems, and Leaves | 771,499,974 | 747,574,430 | 24,927 | | |
| | Illumina HiSeq platform 4000 | Trinity 2.4.0 | Leaves | – | 269,267,462 | 60,597 | | |
| | Illumina HiSeq 2500 platform | Trinity | Roots (R) | R | 54,469,054, 71,462,678 | 54,433,348, 71,35,890 | 299,107 | Shen et al., 2017 |
| | | Stems (S) | S | 50,076,260, 64,920,086 | 50,076,260, 64,826,004 | | | |
| | | Leaves (L) | L | 73,647,052, 53,904,216 | 73,534,024, 53,862,708 | | | |
| | | Flowers (F) | F | 38,776,952, 38,669,310 | 38,736,660, 38,602,508 | | | |
| 454 GS FLX Titanium platform | | | Stems | 553,084 | 518,223 | 36,407 | Guo et al., 2013 |
| | Illumina HiSeq™ 2000 | Trinity | Leaves and Pseudobulbs | 568,756,484 | 583,154,602 | 72,797 | Zhang et al., 2021 |
| Gastrodia elata hybrid (Gastrodia elata Bl.f.elata × Gastrodia elata Bl.f.pilifera) | Illumina HiSeq™ 2000 | | Tuber | 20,611,556 | 20,237,474 | 34,323 | Wang et al., 2020 |
| | | | | | | | |
| Gastrodia elata | BGISEQ-500 platform | Trinity | tuber, stem and flowers | – | – | 113,067 | Shan et al., 2021 |
| | 454 and Solexa, Sanger sequencing | | | – | – | 121,917 | Sedeeek et al., 2013 |
| Ophrys exaltata, O. sibogodes and O. garganica | | | | | | | |
| Paphiopedilum armeniacum | Illumina HiSeq™ 2000 | Trinity v2.4.0 | Capsules | – | – | 183,737 | Fang et al., 2020 |
| Paphiopedilum hirsutissimum | Illumina HiSeq™ 2000 | Trinity (version: v2.9.0) | Flowers | – | 18,236,750–21,697,775 | 28,805–34,806 | Li et al., 2021b |
| Phalaenopsis amabilis white cultivar (Baiyuzan) | Illumina HiSeq™ 2500 platform | Trinity | Petals of White (WP) | WP | 50,282,202 | 19,744,124 | 114,293 | Meng et al., 2019 |
| | | | cultivar | 47,998,340 | 28,738,568 | | | |
| | | | Petals of Purple (PP) | PP | 49,944,218 | 49,091,862 | | | |
| | | | cultivar | 50,589,170 | 49,558,168 | | | |
| | | | 41,232,748 | 40,475,996 | | | |
| Phalaenopsis amabilis | Illumina HiSeq™ 2000 system | Trinity | Petals (P) | P | 10,734,813 | 37,723 | Yang et al., 2014 |
| | | Labellum (L) | L | 16,224,038 | 34,020 | | | |
the case of *Dendrobium huoshanense* where the SwissProt similarity of unigenes was the highest (Zhou et al., 2020). Out of 186,865 unigenes identified in *Anoectochilus roxburghii*, approximately 35, 32, and 47% were annotated using KEGG, SwissProt, and Nr database (Chen et al., 2020). However, only 9,946 out of 73,528 unigenes were annotated by KEGG in *Calanthe tsoongiana* (Jiang et al., 2021a). In *Dendrobium officinale*, the unigenes characterized using SwissProt varied from 12,877 genes (Chen et al., 2019) to 62,695 unigenes (Wang et al., 2021). The variation could be due to the use of different platforms used for sequencing or assembly and due to the type of tissue used in different studies.

### TABLE 1 | Continued

| Plant name                  | Sequencing platform | Assembly     | Plant part | Raw reads | Final reads | Total unigenes | References    |
|-----------------------------|---------------------|--------------|------------|-----------|-------------|----------------|---------------|
| Red *Phalaenopsis* Dtps. Jiuhbao Red Rose | Illumina HiSeq™ 2000 | Trinity software (version trinymaseq,r2012-03-17) | Bud (RB) | –         | RB          | 8,889,080       | 51,771        | Gao et al., 2016 |
| Yellow *Phalaenopsis* Dtps. Fuller’s Sunset |                  |              |            |           |             |                |               |               |
| *Phalaenopsis* hybrid: Konggangjini | Illumina HiSeq2000 | Trinity | Leaf       | 118,996,000 | 79,434,350  | 21,348 genes   | 31,708 isoforms | 80,525        | Xu et al., 2015  |
| *Pleione limprichtii* | Illumina HiSeq™ 4,000 | Trinity | Flower petals and Lips | – | – | 4,955,918 | – | Mohd-Hairul et al., 2020 |
| *Vanilla planifolia* | MiSeq Desktop Sequencer (Illumina) | CLC Genomic Workbench software (Version 6.0) | Tepals | 4,955,918 | 4,826,959 | 1,678,293 | 301,459 contigs | Rao et al., 2014  |

### TABLE 2 | Functional Annotation using KEGG, SwissProt, and non-redundant (Nr) database.

| Plant name | KEGG | SwissProt | Nr database | References    |
|------------|------|-----------|-------------|---------------|
| *Anoectochilus roxburghii* | 66,542 unigenes | 59,736 unigenes | 87,781 unigenes | Chen et al., 2020 |
| *Calanthe tsoongiana* | 9,946 unigenes | 25,124 unigenes | 35,368 unigenes | Jiang et al., 2021a |
| *Cymbidium goeringii* | 33,417 unigenes | 36,911 unigenes | 54,640 unigenes | Ramya et al., 2019 |
| *Cymbidium longibracteatum* | 10,723 unigenes | 21,297 unigenes | 33,487 unigenes | Jiang et al., 2018 |
| *Cymbidium tortisepalum* var. *longibracteatum* | 44,141 | 44,577 | 70,576 | Jiang et al., 2021b |
| *Dactylorhiza hatagirea* | 9,130 transcripts | – | – | Dhiman et al., 2019 |
| *Dendrobium catenatum* | 4,203 unigenes | – | – | Lei et al., 2018 |
| *Dendrobium huoshanense* | 112,603 unigenes | 225,288 unigenes | 140,919 unigenes | Zhou et al., 2020 |
| *Dendrobium nobile* | 108,417 unigenes | 101,132 unigenes | 196,739 unigenes | Yuan et al., 2018 |
| *Dendrobium officinale* | 18,911 unigenes | 48,431 unigenes | 56,378 unigenes | Li et al., 2017 |
| *Gastrodia elata* hybrid (Gastrodia elata *Bl.f.elata* × Gastrodia elata *Bl.f.pilifera*) | 8,364 unigenes | 19,028 unigenes | 24,230 unigenes | Wang et al., 2020 |
| *Gastrodia elata* | 56,585 | 52,164 | 71,069 | Shan et al., 2021 |
| *Ophrys exaltata, O. sphegodes and O. garganica* | 7,394 transcripts | – | – | Sedeek et al., 2013 |
| *Paphiopedilum amnicum* | 12,141 unigenes | 44,893 unigenes | 89,289 unigenes | Fang et al., 2020 |
| *Phalaenopsis amabilis white cultivar (Ba'iyuzuan)* | 16,777 unigenes | – | 48,071 unigenes | Meng et al., 2019 |
| *Red Phalaenopsis Yellow Phalaenopsis* hybrid: Konggangjini | 5,446 unigenes | 19,446 unigenes | 27,084 unigenes | Gao et al., 2016 |
| *Pleione limprichtii* | 14,099 unigenes | – | – | Xu et al., 2015 |
| *Vanilla planifolia* | 11,067 unigenes | 21,177 unigenes | 33,459 unigenes | Zhang et al., 2020b |

In *Phalaenopsis* hybrid, the use of different platforms and tissue types could also contribute to the variation in the number of annotated unigenes.
The annotation of genes or transcripts obtained using various servers helped in the characterization of genes based on their functional roles. The KEGG analysis of different studies in association with pathways of secondary metabolism has been summarized in Table 3. KEGG analysis of stem, leaves, and roots revealed the presence of cyanoamino acid metabolism, phenylpropanoid biosynthesis, diterpenoid biosynthesis, flavonoid and flavonol biosynthesis, steroid biosynthesis, and isoflavonoid biosynthesis pathways in A. roxburghii (Chen et al., 2020b). In B. catenatum, related to the MEP pathway through decarboxylation reaction and is regulated by feedback mechanism at both transcriptional and post-translational levels (Hinson et al., 1997) suggested stem-specific accumulation of alkaloids and terpenoids. These secondary metabolites can be grouped into various classes like alkaloids, terpenoids, polyphenols, phenanthrene, benzyln derivatives, etc. Therapeutic effects of different alkaloids especially terpenoid alkaloids have been widely reported in orchids (Sut et al., 2017; Gantait et al., 2021; Ghai et al., 2021). These terpenes alkaloids are formed through the mevalonate (MVA) pathway and methyl-D-erythritol 4-phosphate (MEP) pathway (Figure 2). MVA pathway initiates with acetyl-CoA as a precursor. Acetyl-CoA undergoes a series of catalyization reactions to produce isopentyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). These IPP units are further processed to form sesquiterpenoids. Meanwhile, the MEP pathway begins with the condensation of pyruvate and D-glyceraldehyde-3-phosphate by 1-Deoxy-D-xylulose-5-phosphate (MEP) pathway (Hinson et al., 1997). The isopentyl pyrophosphate (IPP) units are isomerized into dimethylallyl pyrophosphate (DMAPP) by IPP isomerase which is the initiating molecule in terpenoid biosynthesis. The enzyme, 1-deoxy-D-xylulose-5-phosphate synthase (DXS), is key player in controlling influx into the MEP pathway through decarboxylation reaction and is regulated by feedback mechanism through IPP and DMAPP (Banerjee et al., 2013). This can be corroborated by the higher expression of DXS and the terpenoid enzymes in the inflorescences in Arabidopsis (Carretero-Paulet et al., 2002). This process proceeds to form IPP and DMAPP via multistep reactions catalyzed by a series of enzymes. The MEP and MVA pathways are both linked by an intermediary precursor isopentenyl pyrophosphate. Subsequently, the pathways result in the formation of monoterprenoids, diterpenoids, carotenoids, sesquiterpenoids, and some other metabolites. Sesquiterpene alkaloids are the most abundant types of alkaloids of Dendrobium (Chen et al., 2019). Hsiao et al. (2011) reported the identification of 50 unigenes of the MEP and MVA pathways in Phalaenopsis while in Cymbidium goeringii, 32 unigenes of MVA and 38 unigenes of MEP pathway were identified (Ramya et al., 2019). Forty-six unigenes in Dendrobium huoshanense (Yuan et al., 2018) and 36 in Dendrobium officinale (Shen et al., 2017) related to the MEP and MVA pathway were identified. According to Li et al. (2017), isoprene units obtained through the MEP pathway were responsible for the biosynthesis of dendrobine in Dendrobium nobile. The expression of acetyl-CoA acetyltransferase (AACT), mevalonate diphosphate decarboxylase (MVD), phosphomevalonate kinase (PMK), and Alpha-humulene synthase (TPS21) changes upon inoculation of the orchid with MF23 (Mycena sp.) which results in induction of pathway leading to dendrobine biosynthesis (Li et al., 2017). Besides fungal stimulation, methyl jasmionate (MeJA) treatment of D. officinale also results in increased expression of genes associated with MEP and MVA pathway (Chen et al., 2019). Toh et al. (2017) also studied the fragrant sites in Vanda Mimi Palmer which indirectly points toward the sites of high monoterpenoid production. Higher expression of 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) and 1-deoxysphingolose-5-phosphate synthetase (DXS) was observed in root than in leaf but DXS and 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR) were abundant mainly in stems of Dendrobium huoshanense (Yuan et al., 2018). Yuan et al. (2018) suggested stem-specific accumulation of alkaloids in D. huoshanense but leaf-specific accumulation is observed in D. officinale (Shen et al., 2017). A series of enzymes associated with strictosidine were also identified in a study on Dendrobium officinale (Shen et al., 2017). Accumulation of dendrobine, a sesquiterpene alkaloid, was consistently more when the expression of PMK and MVD was high but got reduced as the expression of the aforementioned genes decreased in MF23 infected Dendrobium nobile orchid plant (Li et al., 2017). In the same study, the dendrobine pathway was negatively correlated with the expression of TPS21 but no relation with genes of MEP was observed (Li et al., 2017). Secologanin synthase (SCS) which is essential for the synthesis of secologanin has also been reported to be involved in alkaloid biosynthesis (Guo et al., 2013). Different terpenes are synthesized from isopentenyl diphosphate through two pathways mevalonate pathway and methylerthythritol phosphate pathway. Hsiao et al. (2006) analyzed transcriptomes of Phalaenopsis bellina and Phalaenopsis equestris where genes related to the DXP-geraniol linalool pathway were identified by data mining. In another study, regulation of monoterpene biosynthesis by PbbHLH4 in Phalaenopsis orchid was provided (Chuang et al., 2018). Terpene synthases (TPSs) are responsible for the structure diversity of terpene while cytochromes P450 (CYPs) further modifies the products from TPSs which provide further diversification of terpenes (Tsai et al., 2017).
## TABLE 3 | KEGG pathway analysis of secondary metabolism.

| Plant name (Reference) | Unigenes/transcripts | Secondary metabolism | Unigenes/transcripts |
|------------------------|----------------------|----------------------|----------------------|
| **Pathway**            | **Unigenes/transcripts** |                      | **Unigenes/transcripts** |
| Anoectochilus roxburghii (Chen et al., 2020) | 66,542 unigenes | Biosynthesis of other secondary metabolites | Root 3,369 unigenes |
|                        |                      |                      | Stem 3,302 unigenes   |
|                        |                      |                      | Leaf 3,280 unigenes   |
| Calanthe tsongiana (Jiang et al., 2021a) | 9,946 unigenes in 25 pathways | Anthocyanin biosynthesis 9 unigenes |  |
|                        |                      | Indole alkaloid biosynthesis 21 unigenes |  |
|                        |                      | Isoflavonoid biosynthesis 36 unigenes |  |
|                        |                      | Tropane, piperidine and pyridine alkaloid biosynthesis 50 unigenes |  |
|                        |                      | Isoquinoline alkaloid biosynthesis 51 unigenes |  |
|                        |                      | Monoterpenoid biosynthesis 56 unigenes |  |
|                        |                      | Sesquiterpenoid and triterpenoid biosynthesis 75 unigenes |  |
|                        |                      | Flavone and flavonol biosynthesis 134 unigenes |  |
|                        |                      | Diterpenoid biosynthesis 172 unigenes |  |
|                        |                      | Terpenoid backbone biosynthesis 197 unigenes |  |
|                        |                      | Flavonoid biosynthesis 236 unigenes |  |
|                        |                      | Phenylpropanoid biosynthesis 466 unigenes |  |
|                        |                      | Biosynthesis of secondary metabolites 3,197 unigenes |  |
|                        |                      | Flavonoid biosynthesis 31 unigenes |  |
| Calanthe tsoongiana | 9,946 unigenes in 25 pathways | Biosynthesis of other secondary metabolites | 290 unigenes |
| Cymbidium goeringii (Ramya et al., 2019) | 33,417 unigenes | Anthocyanin biosynthesis 9 unigenes |  |
|                        |                      | Indole alkaloid biosynthesis 21 unigenes |  |
|                        |                      | Isoflavonoid biosynthesis 36 unigenes |  |
|                        |                      | Tropane, piperidine and pyridine alkaloid biosynthesis 50 unigenes |  |
|                        |                      | Isoquinoline alkaloid biosynthesis 51 unigenes |  |
|                        |                      | Monoterpenoid biosynthesis 56 unigenes |  |
|                        |                      | Sesquiterpenoid and triterpenoid biosynthesis 75 unigenes |  |
|                        |                      | Flavone and flavonol biosynthesis 134 unigenes |  |
|                        |                      | Diterpenoid biosynthesis 172 unigenes |  |
|                        |                      | Terpenoid backbone biosynthesis 197 unigenes |  |
|                        |                      | Flavonoid biosynthesis 236 unigenes |  |
|                        |                      | Phenylpropanoid biosynthesis 466 unigenes |  |
|                        |                      | Biosynthesis of secondary metabolites 3,197 unigenes |  |
| Dendrobium catenatum (Lai et al., 2018) | 4,203 unigenes | Biosynthesis of other secondary metabolites | 2,237 unigenes |
| Dendrobium huoshanense (Zhou et al., 2020) | 112,803 unigenes annotated in 131 pathways | Phenylpropanoid biosynthesis 92 unigenes |  |
|                        | 108,417 unigenes annotated to 33 pathways | Flavonoid biosynthesis 39 unigenes |  |
|                        | 18,911 unigenes assigned to 131 pathways | Flavone and flavonol biosynthesis 18 unigenes |  |
|                        | 12,877 genes grouped into 19 secondary level pathways | Anthocyanin biosynthesis 7 genes |  |
|                        | 20,274 unigenes | Biosynthesis of other secondary metabolites | 5 unigenes |
| Dendrobium officinale (Chen et al., 2019) | 8,364 unigenes | Phenylpropanoid biosynthesis 92 unigenes |  |
|                        |                      | Flavonoid biosynthesis 39 unigenes |  |
|                        |                      | Flavone and flavonol biosynthesis 18 unigenes |  |
|                        |                      | Tropane, piperidine and pyridine alkaloid biosynthesis 13 unigenes |  |
|                        |                      | Isoquinoline alkaloid biosynthesis 8 unigenes |  |
|                        |                      | Anthocyanin biosynthesis 1 unigene |  |
|                        |                      | Biosynthesis of other secondary metabolites | 252 transcripts |
| Dendrobium officinale (Guo et al., 2013) | 7,394 transcripts | Biosynthesis of other secondary metabolites | 5 unigenes |
| Gastrodia elata hybrid (Gastrodia elata Bl.f.elata ×Gastrodia elata Bl.f.pilifera) (Wang et al., 2020) | 16,777 unigenes assigned to 129 pathways | Phenylpropanoid synthesis 168 genes |  |
|                        | 16,777 unigenes assigned to 129 pathways | Flavonoid synthesis 39 genes |  |
|                        | 16,777 unigenes assigned to 129 pathways | Flavone and flavonol synthesis 19 genes |  |
|                        | 16,777 unigenes assigned to 129 pathways | Anthocyanin synthesis 7 genes |  |
|                        | 16,777 unigenes assigned to 129 pathways | Biosynthesis of other secondary metabolites | 328 genes |
| Gastrodia elata hybrid (Gastrodia elata Bl.f.elata ×Gastrodia elata Bl.f.pilifera) (Wang et al., 2020) | 8,364 unigenes | Phenylpropanoid biosynthesis 92 unigenes |  |
|                        | 8,364 unigenes | Flavonoid biosynthesis 39 unigenes |  |
|                        | 8,364 unigenes | Flavone and flavonol biosynthesis 18 unigenes |  |
|                        | 8,364 unigenes | Anthocyanin biosynthesis 1 unigene |  |
|                        | 8,364 unigenes | Biosynthesis of other secondary metabolites | 252 transcripts |
| Ophrys exaltata | 7,394 transcripts | Flavone and Flavonol biosynthesis 13 genes |  |
| O. sphegodes | 7,394 transcripts | Isoquinoline alkaloid biosynthesis 11 genes |  |
| O. garganica (Sadeek et al., 2013) | 7,394 transcripts | Phenylpropanoid biosynthesis 49 genes |  |
| Ophrys exaltata | 7,394 transcripts | Flavonoid biosynthesis 21 genes |  |
| O. sphegodes | 7,394 transcripts | Anthocyanin biosynthesis 1 gene |  |
| O. garganica (Sadeek et al., 2013) | 7,394 transcripts | Flavone and Flavonol biosynthesis 13 genes |  |
| Phalaenopsis amabilis white cultivar (Baiyuzan) | 5,446 unigenes | Isoquinoline alkaloid biosynthesis 11 genes |  |
| Phalaenopsis amabilis purple cultivar (Baolonghuanghou) (Meng et al., 2019) | 5,446 unigenes | Phenylpropanoid biosynthesis 49 genes |  |
| Phalaenopsis amabilis white cultivar (Baiyuzan) | 5,446 unigenes | Flavonoid biosynthesis 21 genes |  |
| Phalaenopsis amabilis purple cultivar (Baolonghuanghou) (Meng et al., 2019) | 5,446 unigenes | Anthocyanin biosynthesis 1 gene |  |
| Phalaenopsis amabilis white cultivar (Baiyuzan) | 5,446 unigenes | Flavone and Flavonol biosynthesis 13 genes |  |
| Phalaenopsis amabilis purple cultivar (Baolonghuanghou) (Meng et al., 2019) | 5,446 unigenes | Isoquinoline alkaloid biosynthesis 11 genes |  |
(Continued)
**TABLE 3 | Continued**

| Plant name (Reference) | Unigenes/transcripts | Secondary metabolism |
|------------------------|----------------------|----------------------|
| Phalaenopsis hybrid: Konggangjinti (Xu et al., 2015) | 14,099 unigenes assigned to 123 pathways | Biosynthesis of secondary metabolites 791 unigenes Terpenoid backbone biosynthesis 55 unigenes Indole alkaloid biosynthesis 1 unigene Monoterpenoid biosynthesis 1 unigene Diterpenoid biosynthesis 20 unigenes Sesquiterpenoid and triterpenoid biosynthesis 4 unigenes Phenylpropanoid biosynthesis 75 unigenes Flavonoid biosynthesis 34 unigenes Flavone and flavonol biosynthesis 15 unigenes Isoquoinoline alkaloid biosynthesis 9 unigenes Tropane, piperidine and pyridine alkaloid biosynthesis 20 unigenes |
| Pleione limprichtii (Zhang et al, 2020b) | 11,067 unigenes mapped onto 131 pathways | Biosynthesis of secondary metabolites 1,294 unigenes Phenylpropanoid biosynthesis 167 unigenes Terpenoid backbone biosynthesis 53 unigenes Flavonoid biosynthesis 36 unigenes Diterpenoid biosynthesis 35 unigenes Isoquoinoline alkaloid biosynthesis 28 unigenes Tropane, piperidine and pyridine alkaloid biosynthesis 24 unigenes Flavone and flavonol biosynthesis 7 unigenes Sesquiterpenoid and triterpenoid biosynthesis 5 unigenes Anthocyanin biosynthesis 1 unigene |

*D. huoshanense*, 229 unigenes of the P450 superfamily were identified (Yuan et al., 2018) but in *D. officinale*, 236 unigenes associated with P450 were mined (Shen et al., 2017). Strictosidine synthase had higher expression levels in protocorm like bodies (PLBs) than in leaves suggesting the higher content of total alkaloid is related to the higher amount of precursor strictosidine produced in *D. officinale* (Wang et al., 2021). The positive stimulation of either MEP or MVA pathway could eventually lead to an increase in the production of alkaloids which could eventually increase the therapeutic potential of the orchid plant.

Besides alkaloids, the role of flavonoids as antioxidant, anti-cancer, and anti-aging agents has also been highlighted (Middleton et al., 2000). The flavonoids are compounds with bridged phenyl rings which are synthesized through the phenylpropanoid pathway. Flavonoid also provides resistance against disease and insects in plants and enable the plant for adapting to adverse environmental conditions with the help of increased production in secondary metabolites (Campos and Hamm, 2000; Yuan et al., 2020). *Anoectochilus roxburghii* is rich in flavonoid compounds, such as dihydroquercetin, quercetin, kaempferol, and myricetin (Ye et al., 2017), which are responsible for the drug activity of this orchid plant (Chen et al., 2020). Lei et al. (2018) reported about C-glycosides type flavonoids are more abundant than O-glycosides in *Dendrobium*. The metabolic analysis of *Anoectochilus roxburghii* revealed an abundance of flavonoids in leaves than in roots or stems (Chen et al., 2020). The by-products of the shikimate pathway are the precursor for a large assortment of secondary metabolites (Tzn et al., 2012; Takayuki et al., 2013). It is a multistep process that starts with the condensation of phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P; Figure 3). The intermediate chorismite after further processing leads to the independent formation of aromatic amino acids, tryptophan, tyrosine, phenylalanine. Phenylalanine is the precursor for the Phenylpropanoid pathway which ultimately results in the synthesis of flavonoids. PAL is the most important rate limiting fulcram enzyme that links primary metabolism with secondary metabolism (Vogt, 2010; Fraser and Chapple, 2011). A positive correlation between the PAL enzyme activity and accumulation of phenylpropanoid compounds have been widely reported (Bate et al., 1994; Vogt, 2010). Carbon flux into different branches of flavonoid synthesis is regulated by flavonol synthase (FLS; Davies et al., 2003). In *Arabidopsis*, activity of hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT) led to maneuvering of the metabolic flux into flavonoids through Chalcone synthase (CHS) activity (Besseau et al., 2007). Additionally, there are several transcription factors that regulate the gene expression which ultimately controls the metabolic flux. The expression of regulatory molecules like MYB is inversely proportional to lignin production, thus facilitating the metabolic flux toward flavonoid production (Fornale et al., 2010). Similarly, elicitors like salicylic acid and methyl jasmonate positively diverts the metabolic flux toward increased production of secondary metabolites (Creelman and Mullet, 1997; Kessler and Baldwin, 2002). Phenylalanine is catalyzed by phenylalanine ammonia lyase (PAL) to form cinnamate which is converted to p-coumaroyl-CoA by transcinnamate 4-monoxygenase (C4H) and 4-coumaroyl-CoA.
synthase (4CL). *p*-coumaroyl-CoA is further processed by series of different enzymes to form flavonoids, flavonols, flavanones, and anthocyanins.

A total of 15 unigenes encoding seven enzymes of the flavonoid pathway were identified from *D. huoshanense* (Zhou et al., 2020) while 31 and 19 unigenes in *D. catenatum* (Lei et al., 2018) and *Pleione limprichtii* (Zhang et al., 2020b). In a study on *Anoectochilus roxburghii*, inoculation with *Ceratobasidium* sp. AR2 increases the flavonoid content of the plant by upregulating PAL, *chalcone synthase* (*CHS*), 4CL and downregulating of *cinnamate 4-hydroxylase* (*C4H*), and *chalcone isomerase* (*CHI*) genes (Zhang et al., 2020a). In a new cultivar of *Cymbidium longibracteatum* with yellow leaves and tubers, seven unigenes related to flavonoid biosynthesis were upregulated (Jiang et al., 2018). Similarly, expression levels of *CHS*, *CHI*, *dihydroflavonol 4-reductase* (*DFR*), *anthocyanidin synthase*
ANS1, and UDP-glucose: flavonoid-3-O-glucosyltransferase (UFGT) were comparatively higher in red, corroborating with higher anthocyanin content in the red stems of D. candidum (Jia et al., 2021). Similarly, most of the genes involved in anthocyanin biosynthesis were upregulated during floral development of Dendrobium nestor (Cui et al., 2021). Expression of PAL and HMG-CoA reductase was upregulated in the abaxial surface of the tissue of Vanda Mimi Palmer (Toh et al., 2017). The rate of flavonoid production in plants was reported to be controlled by CHS with associated CHI. The higher expression levels of CHS, CHI, flavonol synthase (FLS), DFR, and Anthocyanidin reductase (ANR) in roots than in stems and leaves of A. roxburghii were reported as well (Chen et al., 2020). Upregulation of LAR1, DFR3, flavanone 3-hydroxylase (F3H), CHS1, and CHS2 in leaves facilitates the copious accumulation of flavonoids in leaves of Dendrobium officinale (Yuan et al., 2020). In the same study, Dihydroflavonol reductase (DFR), which is responsible for the conversion of flavonoids into anthocyanin biosynthesis, has higher expression in stems and leaves. MeJA treatment in D. officinale lead to the accumulation of bibenzyl (erianin and gigantol) increased due to upregulation of PAL, 4CL, C4H, and CYP450 (Adejobi et al., 2021). During explant browning in Phalaenopsis sp., higher expression of PhPAL, PhCHS, and Ph4CL was observed which suggest the role of anthocyanin in the early stages of tissue browning (Xu et al., 2015). Similarly, upregulated expression of Pa4CL, PaANS, PaF3H, and PaDFR was detected in purple petal cultivar of Phalaenopsis amabilis (Meng et al., 2019). The study on Phalaenopsis did not identify any DEGs related to CHS, ANS, DFR, and flavonoid-3′-hydroxylase (F3′H) in white petals which could be due to either technical limitations or due to absence of anthocyanin pathway (Yang et al., 2014). Similarly, no transcript of flavonoid-3′,5′-hydroxylase (F3′5′H) was identified from the transcriptome of Ophrys even though 61 transcripts of anthocyanins pathway were mined (Sedeek et al., 2013). Expression of PlCHS, PlCHI, and PlFLS was upregulated in white petals but colored petals had higher expression of PlF3′H, PlDFR, and PlANS in Pleione limprichtii (Zhang et al., 2020b). PAL, 4CL, and C4H were upregulated in 8 and 10 weeks old seeds of Vanilla planifolia (Rao et al., 2014). Expression of trans-resveratrol-di-O-methyltransferase-like (ROMT) encoding gene, responsible for resveratrol biosynthesis, was high in tubers of Dactylorhiza hatagirea (Dhiman et al., 2019). It positively correlates with the fact that tubers of this plant are used as anti-inflammatory, anticarcinogenic, and as a cardioprotective agent. Higher expression of ROMT correlated with the abundant quantity of resveratrol and stilbenes (Dhiman et al., 2019). The role of caffeic acid, coumaric acid, and Caffeoyl...
CoA in the synthesis of resveratrol and stilbenes has also been pointed out in the same study. Genes associated with flavonoid pathways were reported to be regulated by UDP-glycosyltransferase and cytochrome P450 (Liu et al., 2013). DcDT8, a bHLH transcription factor in D. candidum, regulated the anthocyanin production by binding to the promoter region of DcF3′H and DcUFGT (Jia et al., 2021). The above review asserts that transcriptomic approaches can serve as a boon for gene discovery, functional annotation, and expression profiling in non-model organisms.

CONCLUSION

Orchids grow in a variety of habits and habitats mainly owing to the presence of an array of unique secondary metabolites which help these plants sustain the stressful conditions. Therefore, these plants have emerged as important source for bioprospecting following traditional approaches. Omics technology, on the other hand, offer great potential for analysis of the complete metabolic pathways and provides detailed insights to gene function for drug discovery and other therapeutic interventions. The present study is a comprehensive analysis of transcriptomes more than 30 orchids mainly focusing on the alkaloids and flavonoids pathways. It can form the basis of an effective resource for the functional studies on tapping the immense potential of unique orchid secondary metabolites to facilitate development of novel therapeutic products from these plants.

AUTHOR CONTRIBUTIONS

JS conceptualized the work. DG and AK performed the analysis and prepared the original draft. PK, SP, and JS critically reviewed and edited the draft. All the authors have read and approved the final version.

ACKNOWLEDGMENTS

DG is grateful to Council of Scientific and Industrial Research for Senior Research Fellowship (File No. 09/135(0809)/2018-EMR-I). AK is thankful to Department of Science and Technology (DST) for INSPIRE Fellowship for Research Students (File No. DST/INSPIRE/03/2021/002638). JS is thankful for partial financial support received from Department of Science and Technology, Government of India under Promotion of University Research and Scientific Excellence (PURSE) grant scheme. Authors are thankful to Jagdeep Verma for the photographs of orchid plants.

REFERENCES

Adejobi, O. I. I., Guan, J., Yang, L., Hu, J. M., Yu, A., Muraguri, S., et al. (2021). Transcriptomic analyses shed light on critical genes associated with benzyl biosynthesis in Dendrobium officinale. Plan. Theory 10:633. doi: 10.3390/plants10040633
Banerjee, A., Wu, Y., Banerjee, R., Li, Y., Yan, H., and Sharkey, T. D. (2013). Feedback inhibition of deoxy-d-xylulose-5-phosphate synthase regulates the methylenecyclopropane 4-phosphate pathway. J. Biol. Chem. 288, 16926–16936. doi: 10.1074/jbc.M113.464636
Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. doi: 10.1089/cmb.2012.0021
Bao, Q., Qian, L., Gong, C., and Shen, X. (2017). Immune-enhancing activity of polysaccharides from Gastrodia elata. J. Food Process. Preserv. 41:e13016. doi: 10.1111/jfpp.13016
Bate, N. J., Orr, J., Ni, W., Meromi, A., Nadler-Hassar, T., Doerner, P. W., et al. (1994). Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-determining step in natural product synthesis. Proc. Natl. Acad. Sci. U. S. A. 91, 7608–7612. doi: 10.1073/pnas.91.16.7608
Besseau, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B., and Legrand, M. (2007). Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects asexual transport and plant growth. Plant Cell 19, 148–162. doi: 10.1105/tpc.106.044495
Bulpitt, C. J., Li, Y., Bulpitt, P. F., and Wang, J. J. (2007). The use of orchids in Chinese medicine. J. R. Soc. Med. 100, 558–563. doi: 10.1177/01407600710012014
Burkill, I. H. (1935). A Dictionary of the Economic Products of the Malay Peninsula, Vol. II. London: Crown Agents.
Campos, E. G., and Hamdan, F. E. (2000). Cloning of the chaperolin t-complex polypeptide 1 gene from Schistosoma mansoni and studies of its expression levels under heat shock and oxidative stress. Parasitol. Res. 86, 253–258. doi: 10.1007/s004360050039
Carrera, J., Rodrigo, G., Jaramillo, A., and Elena, S. F. (2009). Reverse-engineering the Arabidopsis thaliana transcriptional network under changing environmental conditions. Genome Biol. 10, R96–R15. doi: 10.1186/gb-2009-10-9-96
Carretero-Paulet, L., Ahumada, I., Cunillera, N., Rodriguez-Concepcion, M., Ferrer, A., Boronat, A., et al. (2002). Expression and molecular analysis of the Arabidopsis DXR gene encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase, the first committed enzyme of the 2-C-methyl-D-erythritol 4-phosphate pathway. Plant Physiol. 129, 1581–1591. doi: 10.1104/pp.003798
Chauhan, N. S. (1990). Medicinal orchids of Himachal Pradesh. J. Orchid Soc. India. 4, 99–105.
Chawla, A. S., Sharma, A. K., Handa, S. S., and Dhar, K. L. (1992). Chemical studies and anti-inflammatory activity of Vanda roxburghii roots. Indian J. Pharm. Sci. 54, 159–161.
Chen, Y., Pan, W., Jin, S., and Lin, S. (2020). Combined metabolomic and transcriptomic analysis reveals key candidate genes involved in the regulation of flavonoid accumulation in Anoectochilus roxburghii. Process Biochem. 101, 339–351. doi: 10.1016/j.procbio.2020.01.004
Chen, Y., Wang, Y., Lyu, P., Chen, L., Shen, C., and Sun, C. (2019). Comparative transcriptomic analysis reveal the regulation mechanism underlying MeJA-induced accumulation of alkaloids in Dendrobium officinale. J. Plant Res. 132, 419–429. doi: 10.1007/s10265-019-01099-6
Chen, J., Zhu, F., Liu, L., Yi, L., Dai, Y., Chen, S., et al. (2021). Integrative analyses of transcriptome and metabolome shed light on the regulation of secondary metabolites in pseudobulbs of two Bletilla striata (Thunb.) Reichb. f. varieties. J. Appl. Res. Med. Aromat. Plants 293:110738. doi: 10.1016/j.aprmat.2021.110738
Choukarya, R., Choursia, A., and Rathi, J. (2019). In vivo and in vitro antioxidant activity of hydroalcoholic extract of Dactylorhiza hatagirea roots: an evaluation of possible phytoconstituents. J. Drug Deliv. Ther. 9, 76–81. doi: 10.22270/jddt.v9i6.s3752
Chuang, Y., Huang, H., Tsai, W., Chen, W., and Chen, H. (2018). PbHHL4 regulates floral monoterpane biosynthesis in Phalaenopsis orchids. J. Exp. Bot. 69, 4363–4377. doi: 10.1093/jxb/ery246
Creelman, R. A., and Mullet, J. E. (1997). Biosynthesis and action of jasmonates in Arabidopsis thaliana. Annu. Rev. Plant Biol. 48, 355–381. doi: 10.1146/annurev.arplant.48.1.355
Ghai et al. Comparative Transcriptome Analysis in Orchids... DOI: http://dx.doi.org/10.1007/s12870-021-03288-9

Song, J. L., Kang, J. Y., Yong, H., Kim, Y. C., and Moon, A. (2012). Denbinobin, a phenanthrene from Dendrobium nobile, inhibits invasion and induces apoptosis in SNU-484 human gastric cancer cells. Oncol. Rep. 27, 813–818. DOI: 10.3892/or.2011.1551

Sut, S., Maggi, F., and DallAcqua, S. (2017). Bioactive secondary metabolites from orchids (Orchidaceae). Chem. Biodivers. 14, 1700172. DOI: 10.1002/cbdv.201700172

Takayuki, T., Mutsuimi, W., Rainer, H., and Fernie, A. R. (2013). Shikimate and phenylalanine biosynthesis in the green lineage. Front. Plant Sci. 4:62. DOI: 10.3389/fpls.2013.00062

Teoh, E. S. (2016). Medicinal Orchids of Asia. Cham: Springer.

Teoh, E. S. (2019). Orchids as Aphrodisiac, Medicine or Food. Cham: Springer.

Toh, C., Mohd-Hairul, A. R., Ain, N. M., Nasamivayam, P. G., Ro, G., Abdullah, N. A. P., et al. (2017). Floral morphomicroscopy and transcriptome analyses of a fragrant Vandaceous orchid, Vanda Mimi palmer, for its fragrance production sites. BMC. Res. Notes 10:554. DOI: 10.1186/s13104-017-2872-6

Tsi, W. C., Dievart, A., Hsu, C. C., Hsiao, Y. Y., Chou, S. Y., Huang, H., et al. (2017). Post genomics era for orchid research. Bot. Stud. 22, 1–22. DOI: 10.3390/jims22136947

Tzin, V., Malikys, S., Zvi, M. M. B., Bedair, M., Sumner, L., Aharoni, A., et al. (2012). Expression of a bacterial feedback-inhibitive 3-deoxy-d-arabinoheptulosonate 7-phosphate synthase of the shikimate pathway in Arabidopsis elucidates potential metabolic bottlenecks between primary and secondary metabolism. New Phytol. 194, 430–439. DOI: 10.1111/nph.12052.x

Vij, S. P., Verma, J., and Kumar, S. C. (2013). Orchids of Himalach Pradesh. Blishen Singh Mahendra Pal Singh. Dhradun.
Yang, Y., Wang, J., Ma, Z., Sun, G., and Zhang, C. (2014). De novo sequencing and comparative transcriptome analysis of white petals and red labella in *Phalaenopsis* for discovery of genes related to flower color and floral differentiation. *Acta Soc. Bot. Pol.* 83, 191–199. doi: 10.5586/asbp.2014.023

Ye, S., Zhao, Q., and Zhang, A. (2017). *Anoectochilus roxburghii*: a review of its phytochemistry, pharmacology, and clinical applications. *J. Ethnopharmacol.* 209, 184–202. doi: 10.1016/j.jep.2017.07.032

Yuan, Y., Yu, M., Jia, Z., Song, X., Liang, Y., and Zhang, J. (2018). Analysis of *Dendrobium huoshanense* transcriptome unveils putative genes associated with active ingredients synthesis. *BMC Genomics* 19:978. doi: 10.1186/s12864-018-5305-6

Zhebino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

Zhang, Y., Cai, J., Ruan, H., Pi, H., and Wu, J. (2007). Antihyperglycemic activity of kinsenoside, a high yielding constituent from *Anoectochilus roxburghii* in streptozotocin diabetic rats. *J. Ethnopharmacol.* 114, 141–145. doi: 10.1016/j.jep.2007.05.022

Zhang, C., Chen, J., Huang, W., Song, X., and Niu, J. (2021a). Transcriptomics and metabolomics reveal purine and phenylpropanoid metabolism response to drought stress in *Dendrobium sinense*, an endemic orchid species in Hainan Island. *Front. Physiol.* 12:692702. doi: 10.3389/fphys.2021.692702

Zhang, Y., Li, Y., Chen, X., Meng, Z., and Guo, S. (2020a). Combined metabolome and transcriptome analyses reveal the effects of mycorrhizal fungus *Ceratobasidium* sp. AR2 on the flavonoid accumulation in *Anoectochilus roxburghii* during different growth stages. *Int. J. Mol. Sci.* 21:564. doi: 10.3390/ijms21020564

Zhang, M., Yu, Z., Zeng, D., Si, C., Zhao, C., Wang, H., et al. (2021b). Transcriptome and metabolome reveal salt-stress responses of leaf tissues from *Dendrobium officinale*. *Biomol. Ther.* 11:736. doi: 10.3390/biom11050736

Zhang, Y., Zhang, L., Liu, J., Liang, J., Si, J., and Wu, S. (2017). *Dendrobium officinale* leaves as a new antioxidant source. *J. Funct. Foods* 37, 400–415. doi: 10.1016/j.jff.2017.08.006

Zhang, Y., Zhou, T., Dai, Z., Dai, X., Li, W., Cao, M., et al. (2020b). Comparative transcriptomics provides insight into floral color polymorphism in a *Pleione limprichtii* orchid population. *Int. J. Mol. Sci.* 21:247, doi: 10.3390/ijms21010247

Zhang, C., Liu, Q., Halaweish, F., Shao, B., Ye, Y., and Zhao, W. (2003). Copacamphe, picrotoxane, and alloaromadendrane sesquiterpene glycosides and phenolic glycosides from *Dendrobium moniliforme*. *J. Nat. Prod.* 66, 1140–1143. doi: 10.1021/np0301801

Zhou, P., Pu, T., Gui, C., Zhang, X., and Gong, L. (2020). Transcriptome analysis reveals biosynthesis of important bioactive constituents and mechanism of stem formation of *Dendrobium huoshanense*. *Sci. Rep.* 10:2857. doi: 10.1038/s41598-020-59737-2

Zhou, Z., Ying, Z., Wu, Z., Yang, Y., Fu, S., Xu, W., et al. (2021). Anthocyanin genes involved in the flower coloration mechanisms of *cymbidium kanran*. *Front. Plant Sci.* 12:737815. doi: 10.3389/fpls.2021.737815

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Ghai, Kaur, Kahlon, Pawar and Sembi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**