From Plant to Yeast—Advances in Biosynthesis of Artemisinin

Le Zhao 1,2, Yunhao Zhu 1,2, Haoyu Jia 1, Yongguang Han 1, Xiaoke Zheng 1,2, Min Wang 3,4,* and Weisheng Feng 1,2,*

1 School of Pharmacy, Henan University of Chinese Medicine, Zhengzhou 450046, China
2 Co-Construction Collaborative Innovation Center for Chinese Medicine and Respiratory Diseases by Henan and Education Ministry of P. R. China, Henan University of Chinese Medicine, Zhengzhou 450046, China
3 College of Chemistry and Materials Engineering, Beijing Technology and Business University, Beijing 100048, China
4 Beijing Key Laboratory of Plant Research and Development, Beijing Technology and Business University, Beijing 100048, China
* Correspondence: wangmmin_216@163.com (M.W.); fwsh@hactcm.edu.cn (W.F.); Tel.: +86-13426292115 (M.W.); +86-371-60190296 (W.F.)

Abstract: Malaria is a life-threatening disease. Artemisinin-based combination therapy (ACT) is the preferred choice for malaria treatment recommended by the World Health Organization. At present, the main source of artemisinin is extracted from Artemisia annua; however, the artemisinin content in A. annua is only 0.1–1%, which cannot meet global demand. Meanwhile, the chemical synthesis of artemisinin has disadvantages such as complicated steps, high cost and low yield. Therefore, the application of the synthetic biology approach to produce artemisinin in vivo has magnificent prospects. In this review, the biosynthesis pathway of artemisinin was summarized. Then we discussed the advances in the heterologous biosynthesis of artemisinin using microorganisms (Escherichia coli and Saccharomyces cerevisiae) as chassis cells. With yeast as the cell factory, the production of artemisinin was transferred from plant to yeast. Through the optimization of the fermentation process, the yield of artemisinic acid reached 25 g/L, thereby producing the semi-synthesis of artemisinin. Moreover, we reviewed the genetic engineering in A. annua to improve the artemisinin content, which included overexpressing artemisinin biosynthesis pathway genes, blocking key genes in competitive pathways, and regulating the expression of transcription factors related to artemisinin biosynthesis. Finally, the research progress of artemisinin production in other plants (Nicotiana, Physcomitrella, etc.) was discussed. The current advances in artemisinin biosynthesis may help lay the foundation for the remarkable up-regulation of artemisinin production in A. annua through gene editing or molecular design breeding in the future.

Keywords: Artemisia annua; artemisinin; synthetic biology; genetic engineering; transcription factor

1. Introduction

Plants fix carbon through photosynthesis and use solar energy to convert water and CO₂ into organic matter to produce primary metabolites, which are used as precursors to synthesize a series of secondary metabolites through secondary metabolism. These secondary products are not only important for plant adaptation and response to the environment, but are also the source of many clinical drugs [1]. According to the chemical structure and biosynthesis pathway, plant secondary metabolites can be divided into three major groups: terpenoids (54%), alkaloids (27%), phenolics (18%) and others (1%) [2]. Terpenoids are the largest number of plant secondary products, possessing diversified structures and functions. Based on the number of 5-carbon isoprene units, terpenoids are classified into hemiterpenoids (5C), monoterpenoids (10C), sesquiterpenoids (15C), diterpenoids (20C), sesterterpenoids (25C), triterpenoids (30C), tetramerpenoids (40C), and polyterpenoids (>40C) [3]. Paclitaxel (taxol), a diterpenoid anticancer drug isolated from Taxus brevifolia, has been approved for the clinical treatment of breast cancer, non-small cell lung
cancer, ovarian cancer and AIDS-related Kaposi sarcoma (used as second-line therapy) [4]. Many plant polyphenolics metabolites, such as flavonoids, phenolic acids, lignans and tannins, are powerful antioxidants. Curcumin, a polyphenol compound that is produced by Curcuma longa, has antioxidant, anti-inflammatory and anticancer properties [5]. Among other types of plant secondary metabolites, an example is melatonin, which is derived from tryptophan and was initially isolated from animals before being discovered in plants. Melatonin plays an important role in plant growth and development as well as the response to biotic and abiotic stress; in addition, it has strong antioxidant activity [6].

Artemisinin is a plant secondary metabolite of sesquiterpenoids, which was first isolated from the medicinal plant Artemisia annua L. by Tu Youyou in 1972 [7]. Artemisinin-based combination therapy (ACT) is currently the preferred treatment for malaria recommended by the World Health Organization (WHO). Malaria is a life-threatening disease caused by the Plasmodium parasite, which is transmitted to humans through the bite of infected female Anopheles mosquitoes. So far, the global malaria epidemic remains severe. According to the latest World malaria report 2021, malaria infection cases in 2020 were found in 85 countries, infecting more than 241 million people and causing 627,000 deaths, mainly in Africa, Southeast Asia and the Eastern Mediterranean. Compared with 2019, there were 14 million more malaria infections and 47,000 more deaths than in 2020 due to the disruption of medical services during the COVID-19 pandemic [8]. In 2015, Tu Youyou was awarded the Nobel Prize in Physiology and Medicine for her discovery of a new method to treat malaria with artemisinin. Recent studies have found that artemisinin and its derivatives have anti-cancer, antibacterial and immunomodulatory properties—in addition to the efficacy of treating malaria [9]. Moreover, artemisinin can also be used for the treatment of diabetes [10] and pulmonary tuberculosis [11], which have broad application prospects.

The content of artemisinin in A. annua is very low, being only 0.1–1% DW (dry weight) of the plant. The process of extracting artemisinin from A. annua is complicated and costly, which cannot meet the global demand for artemisinin [12]. Although the chemical synthesis of artemisinin was successful as early as 1983, it has not been put into commercial production due to the disadvantages of having multiple steps, a high cost and low yield [13]. Therefore, how to obtain artemisinin efficiently, economically and in large quantities has become an international research hotspot. With the development of molecular biology, the application of synthetic biology to produce active components of medicinal plants has broad application prospects [14]. In this review, the artemisinin biosynthesis pathway was summarized, then reconstituting the biosynthesis pathway of artemisinin using microorganisms as chassis cells was discussed, and finally, the research progress in the production of artemisinin from different plants was reviewed, which laid the foundation for the mass, rapid and low-cost acquisition of artemisinin.

2. The Artemisinin Biosynthesis Pathway

In order to generate artemisinin in microorganisms and plants with the synthetic biology method, the biosynthesis pathway of artemisinin needs to be dissected. Artemisinin was derived from the plant terpenoid biosynthesis pathway by using IPP (isopentenyl diphosphate) and its isomer, DMAPP (dimethylallyl diphosphate), as raw materials, which were produced by the MVA (mevalonate) pathway in the cytoplasm and the MEP (methylerthritol phosphate) pathway in the plastid. Derivatives of the MVA pathway were precursors to sesquiterpenes, triterpenes andpolyterpenoids, while the MEP pathway was generally involved in the biosynthesis of monoterpens, diterpenes and tetraterpenes [15]. Two molecules of IPP and one molecule of DMAPP were catalyzed by FPPS (farnesyl diphosphate synthase) to form 15-carbon FPP (farnesyl diphosphate), which entered the downstream metabolic pathway of artemisinin biosynthesis (Figure 1).

ADS (Amorpha-4,11-diene synthase) catalyzed the cyclization of FPP to form amorphine-4,11-diene, the precursor of artemisinin [16]. Under the coordination synergistic effect of CYP71AV1 (cytochrome P450 monooxygenase) and CPR (cytochrome P450
reductase), amorpha-4,11-diene was further oxidized to artemisinic alcohol and artemisinic aldehyde [17,18]. Artemisinic aldehyde was catalyzed by CYP71AV1 and ALDH1 (aldehyde dehydrogenase 1) to generate artemisinic acid [19]. Starting from artemisinic aldehyde, an additional branch emerged. Artemisinic aldehyde was catalyzed by DBR2 (artemisinic aldehyde Δ11(13) reductase) to generate DHAAA (dihydroartemisinic aldehyde), and DHAAA was converted to DHAA (dihydroartemisinic acid) by ALDH1 [20]. Artemisinic acid and DHAA were thought to be the final compounds formed by the enzyme-catalyzed reactions in this pathway (Figure 1). In the subcuticular space of glandular-secreting trichomes (GSTs) on the surface of *A. annua* leaves, artemisinic acid was converted to artemisinin B and DHAA to artemisinin through photo-oxidation, which are a series of non-enzymatic reactions [21,22].

At present, artemisinin has been produced by the synthetic biology approach, mostly through the semi-synthetic route. The precursors of artemisinin, such as amorpha-4,11-diene, artemisinic acid and DHAA, were prepared in microorganisms by metabolic engineering, and then artemisinin was generated by chemical synthesis; however, some studies suggested that there may be a specifically expressed and highly active peroxidase in the GSTs of *A. annua*, which can catalyze the conversion of DHAA to artemisinin [23]. If such a peroxidase existed, the total biosynthesis of artemisinin would be carried out in microorganisms, rather than semi-synthesis.

**Figure 1.** Artemisinin biosynthesis pathway and competitive metabolic pathway in *A. annua*. The biosynthesis of artemisinin is derived from the MVA and MEP pathways. FPP is the important intermediate at the branch point of artemisinin biosynthesis, which can not only be converted into amorpha-4,11-diene and artemisinin, but also into other terpenoids through competitive metabolic pathways, such as squalene, β-farnesene, β-caryophyllene, germacrene A, etc. The competitive
3. Metabolic Engineering in Microorganisms

With the advancement of sequencing technology, more and more medicinal plant genomes have been sequenced, and the biosynthetic pathways of many natural products have been elucidated [24]. On this basis, the biosynthetic pathway of natural products can be reconstructed in microorganisms, and natural products and their precursors are able to be obtained rapidly and in large quantities through biological fermentation, thereby laying a foundation for the commercial application of natural plant products. As an effective antimalarial drug, the development of artemisinin was restricted by the cost of extracting compounds from natural A. annua and the complexity of chemical synthesis; however, for those who are most affected by malaria (such as African areas), classical plant breeding may not be enough to reduce the production cost of artemisinin to an affordable price. Therefore, using microorganisms (Escherichia coli and Saccharomyces cerevisiae) as chassis cells, the heterologous biosynthesis of artemisinin or its precursor through metabolic engineering may provide an alternative source of artemisinin supply. It took Keasling’s laboratory ten years to achieve the semi-synthesis of artemisinin, thereby transferring the production of artemisinin from plants to microorganisms.

In 2003, Martin et al. transformed the codon-optimized ADS gene into E. coli, simultaneously co-expressed the E. coli SOE4 operon (encoding the rate-limiting enzyme genes dxs, ippHp and ispA of the E. coli MEP pathway), and heterologously expressed the MVA pathway genes (pMevT and pMBIS plasmids) from S. cerevisiae. This was the first time that amorpha-4,11-diene, the precursor of artemisinin, was synthesized in E. coli at a concentration of 24 mg/L [25] (Table 1). After optimization of the fermentation conditions, the yield of amorpha-4,11-diene increased 20-fold to 0.5 g/L [26]. Tsuruta et al. continued to optimize the genes of the yeast MVA pathway for heterologous expression in E. coli and replaced the yeast HMGS and THMG1 genes on the pMevT plasmid with more active equivalent genes from Staphylococcus aureus. Combined with the restricted supply of carbon and nitrogen, the amorpha-4,11-diene production in E. coli reached 27.4 g/L [27]. On the basis of the heterologous expression of the yeast MVA pathway genes (pAM92 plasmid), the ADS, AMO (amorpha-4,11-diene oxidase) and CPR genes of A. annua were simultaneously expressed to produce artemisinic acid in E. coli with the highest yield being 105 mg/L [28].

<table>
<thead>
<tr>
<th>Enzyme Name</th>
<th>Expression Mode</th>
<th>Gene Source</th>
<th>Host</th>
<th>Products</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMevT (atoB; tHMGS, tHMGR)</td>
<td>Overexpression</td>
<td>Yeast</td>
<td>E. coli</td>
<td>Amorpha-4,11-diene</td>
<td>[25]</td>
</tr>
<tr>
<td>pMBIS (EGR12, EGR8, MVD1, idi; ispA)</td>
<td>Overexpression</td>
<td>Yeast</td>
<td>E. coli</td>
<td>Amorpha-4,11-diene</td>
<td>24 mg/L</td>
</tr>
<tr>
<td>pSOE4 (dxs, ippHP; ispA)</td>
<td>Overexpression</td>
<td>Yeast</td>
<td>A. annua</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Metabolic engineering to enhance the production of amorpha-4,11-diene and artemisinic acid in E. coli or yeast.

pathways are marked with red boxes. HMG-CoA, 3-hydroxy-3-methylglutaryl coenzyme A; GA-3P, glyceraldehyde 3-phosphate; DXP, 1-deoxy-D-xylulose 5-phosphate; MEPP, 2-C-methyl-D-erythritol-4-phosphate; HMBPP, 4-hydroxy-3-methyl-but-2-enyl diphosphate; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase; DXS, 1-deoxy-D-xylulose 5-phosphate synthase; DXR, 1-deoxy-D-xylulose 5-phosphate reductoisomerase; HDR, 4-hydroxy-3-methyl-but-2-enyl diphosphate reductase; IPP1, isopentenyl diphosphate isomerase; FPPS, farnesyl diphosphate synthase; ADS, amorpha-4,11-diene synthase; CYP71AV1, cytochrome P450 monoxygenase; CPR, cytochrome P450 reductase; ALDH1, aldehyde dehydrogenase 1; DRB2, artemisinic aldehyde Δ11(13) reductase; SQS, squalene synthase; BFS, β-farnesene synthase; CPS, β-caryophyllene synthase; GAS, germacrene A synthase; RED1, dihydroartemisinic aldehyde reductase.
| pAM45 (atoB1, HMGS, tHMGR, MK, PMK, PMD, id12, ispA3) | Overexpression | Yeast | *E. coli* | Amorpha-4,11-diene 293 mg/L | [29] |
| pAM94 (ADS5, MK) | | | | | |
| pAM52 (atoB1, mvaS, mvaA)5 | Overexpression | Yeast | *E. coli* | Amorpha-4,11-diene 27.4 g/L | [27] |
| pMBIS (EGR12, EGR8, MVD1, id12, ispA3) | Overexpression | *A. annua* | | | |
| pAM92 (atoB1, HMGS, tHMGR, EGR12, EGR19, id12, ispA3, ADS5) | Overexpression | Yeast | *E. coli* | Artemisinic acid 105 mg/L | [28] |
| AMO, CPR | | | | | |
| MVA pathway enzymes (EGR13, tHMGR, EGR12, EGR8, EGR20) | Overexpression | Yeast | Yeast | Artemisinic acid 100 mg/L | [18] |
| ADS, CYP71AV1, CPR | Down-regulation | Yeast (CEN.PK2) | | Amorpha-4,11-diene 40 g/L | [30] |
| All enzymes of MVA pathway (EGR10, EGR13, tHMG1, EGR12, EGR8, EGR19, IDI1, EGR20) | Overexpression | Yeast | Yeast (CEN.PK2) | Amorpha-4,11-diene 40 g/L | [30] |
| EGR9 | Down-regulation | Yeast | | | |
| ADS, CYP71AV1, CPR | Overexpression | *A. annua* | | | |
| Every gene of MVA pathway (EGR10, EGR13, tHMG1, EGR12, EGR8, EGR19, IDI1, EGR20) | Overexpression | Yeast | Yeast | Artemisinic acid 25 g/L | [31] |
| EGR9 | Down-regulation | Yeast | | | |
| ADS, CYP71AV1, CPR1, CYB5, ADH1, ALDH1 | Overexpression | *A. annua* | | | |

1 atoB, acetoacetyl-CoA thiolase from *E. coli*. 2 id1, IPP isomerase from *E. coli*. 3 ispA, FPP synthase from *E. coli*. 4 ippHp, IPP isomerase from Haematococcus pluvialis. 5 ADS from *A. annua*. 6 pAM52 was the pMevT derivative, but HMGS and tHMGR were replaced by mvaS and mvaA from *S. aureus*.

In the downstream biosynthetic pathway of artemisinin, the conversion from amorpha-4,11-diene to artemisinic acid and DHAA required the synergistic catalysis of CYP71AV1 and CPR, both of which were associated with the cell’s endomembrane system; however, *E. coli* lacked the CYP450 gene and the endomembrane system of eukaryotic cells, resulting in the affected expression of CYP71AV1 and CPR genes. In 2006, Ro et al. modified the FPP biosynthesis pathway in *S. cerevisiae* to prompt FPP production, which down-regulated the expression of EGR9 and inhibited the conversion of FPP to squalene, thus allowing more FPP to flow into the artemisinic acid biosynthesis pathway. Meanwhile, the *ADS*, CYP71AV1 and CPR genes of *A. annua* were transformed into yeast strain EPY224 (Table 1). Finally, the biosynthesis of artemisinic acid was achieved in yeast with a yield of 100 mg/L [18]. Moreover, the biosynthesized artemisinic acid was able to be secreted outside of the engineered yeast, facilitating the purification of artemisinic acid from the yeast fermentation broth, which further prompted the industrialization process of artemisinin.

Westfall et al. adopted the yeast strain, CEN.PK2, to overexpress all the genes of the MVA pathway while down-regulating the EGR9 gene so that the production of amorpha-4,11-diene reached 40 g/L [30]. In 2013, Paddon et al. transformed the *ADS*, CYP71AV1, CPR1, CYB5, ADH1 and ALDH1 genes of *A. annua* into yeast after codon optimization to reduce the accumulation of artemisinic aldehyde and improved the fermentation process. As a result, the yield of artemisinic acid reached 25 g/L (Table 1), which met the demand of industrial production. Then, artemisinic acid was extracted from the fermentation...
broth, and after four steps of chemical synthesis in vitro, artemisinin was finally acquired with an overall yield of 40–45% [31].

In summary, the use of synthetic biology methods to produce artemisinic acid in microorganisms mainly made the following improvements: 1. Codon optimization. The key genes of artemisinic acid biosynthesis pathway were codon-optimized and transformed into E. coli and yeast. 2. Reconstruction of metabolic pathways and selection of enzymes with high catalytic activity. Enzymes of the yeast MVA pathway were replaced with enzymes from S. aureus. 3. Choice of suitable chassis cells. The CYP450 oxidase and reductase of A. annua were expressed in eukaryotic yeast. 4. Control of the flow of metabolites. In yeast cells, the down-regulation of the ERG9 gene (encoding squalene synthase) inhibited the conversion of FPP to squalene. Meanwhile, the genes of the MVA pathway were overexpressed under strong promoters, thereby allowing more FPP to flow into the artemisinic acid biosynthesis pathway.

4. Production of Artemisinin in Genetically Engineered Plants

4.1. Genetic Engineering in A. annua

The artemisinin content in A. annua can be improved by overexpressing artemisinin biosynthetic genes, blocking key genes in competitive pathways of artemisinin biosynthesis, or regulating the expression of transcription factors involved in artemisinin biosynthesis.

4.1.1. Overexpression of Key Genes in Artemisinin Biosynthesis

In the past two decades, key genes in the artemisinin biosynthesis pathway have been isolated and characterized (Figure 1). The content of artemisinin in A. annua was adjusted by regulating the expression of genes involved in the MVA and MEP pathways, or by overexpressing key genes in the downstream biosynthesis pathway of artemisinin (Table 2).

According to the genetic map and genomic data of A. annua, HMG and DXR genes are closely related to artemisinin content [32,33]. HMG and DXR are the rate-limiting enzymes of the MVA and MEP pathway, which provided the substrates IPP and DMAPP for the biosynthesis of artemisinin. By applying competitive inhibitors (fosmidomycin and mevinolin) to suppress the activities of DXR and HMG, respectively, it was found that the production of artemisinin decreased by 14.2% and 80.4%, indicating that the MVA pathway was the main carbon donor for artemisinin biosynthesis [34]. The HMG gene from Catharanthus roseus was overexpressed in A. annua through Agrobacterium-mediated transformation, and the artemisinin content of transgenic plants reached 0.60 mg/g DW, while artemisinin in non-transgenic lines was only 0.37 mg/g DW [35]. Overexpression of the DXR gene from C. roseus in A. annua prompted a 2.33-fold increase in the yield of artemisinin to 1.21 mg/g DW [36].

Overexpression of the IPP11 and HDR1 genes in A. annua were able to increase the accumulation of artemisinin and artemisinin B, while inhibition of the expression of the HDR1 gene reduced the artemisinin content [37,38]. FPP was the precursor of secondary metabolites such as sesquiterpenes (including artemisinin), triterpenes, coenzyme Q, plastoquinone, etc., so increasing the production of FPP can promote the biosynthesis of artemisinin. The FPPS gene of A. annua was first isolated by Matsushita in 1996 [39], and overexpression of the FPPS gene increased the yield of artemisinin by 1.38–3.36 times [40–42].

ADS catalyzed the conversion of FPP to amorpha-4,11-diene, the first key enzyme in the artemisinin biosynthesis pathway, and the ADS gene was specifically expressed in the GST cells of A. annua [43]. The haplotype genome comparison of A. annua revealed that the copy number of the ADS gene was highly correlated with artemisinin yield [44]. Overexpression of the ADS gene significantly increased not only artemisinin content, but also the content of artemisinic acid and DHAA [45]. DBR2 was another key enzyme in the artemisinin biosynthesis pathway, which was also specifically expressed in the GSTs of A.
annua. DBR2 catalyzed artemisinic aldehyde to form dihydroartemisinic aldehyde, which further generated artemisinin. If the activity of DBR2 decreased, more artemisinic aldehyde was converted into artemisinic acid, resulting in an increased accumulation of artemisinin B [20]. The content of artemisinin in DBR2-overexpressed transgenic plants was 1.50–2.14 mg/g DW, which was 1.59–2.26 times that of the control [46].

The overexpression of a single key gene involved in the artemisinin biosynthesis can increase the artemisinin content, but the increase was not significant. Therefore, simultaneous overexpression of two or more key genes was able to remarkably improve the content of artemisinin [33]. When HMGR and ADS genes were overexpressed in A. annua, the artemisinin content was increased by 8.65-fold to 1.73 mg/g DW [47]. The ADS and FPPS genes were fused together and overexpressed under the regulation of the CaMV3SS or CYP71AV1 promoter. The CaMV3SS promoter significantly up-regulated the expression of the ADS and FPPS genes, but the accumulation of artemisinin did not increase significantly; however, under the control of the CYP71AV1 promoter, the ADS-FPPS fusion gene was specifically expressed in GSTs and significantly enhanced the artemisinin concentration to 26 mg/g DW [48]. Co-expression of the CYP71AV1 and CPR genes can improve the content of artemisinin by 1.38–2.68 times, reaching 0.98–2.44 mg/g DW [36,49]; when the three genes, FPPS, CYP71AV1 and CPR, were co-expressed, the artemisinin content was increased by 3.6 times, reaching 2.98 mg/g FW [50]; when the four genes, ADS, CYP71AV1, CPR and ALDH1, were co-expressed, the artemisinin content reached 27 mg/g DW, which was 3.4 times that of the control [51].

Therefore, the use of transgenic methods to regulate the expression of key genes in the biosynthesis pathway of artemisinin, and the overexpression of one or more key genes, can remarkably improve the content of artemisinin in transgenic plants (Table 2), which was of great significance for obtaining new varieties of A. annua with a high artemisinin content.

### Table 2. Genetic engineering to increase artemisinin production in A. annua.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Expression Mode</th>
<th>MVA or MEP pathway key genes</th>
<th>Artemisinin (mg/g DW)</th>
<th>Fold Change</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Control</td>
<td>Transgenic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CrHMGR</td>
<td>Overexpression</td>
<td>0.32</td>
<td>0.39</td>
<td>1.22</td>
<td>[52]</td>
</tr>
<tr>
<td>CrDXR</td>
<td>Overexpression</td>
<td>0.37</td>
<td>0.60</td>
<td>1.62</td>
<td>[35]</td>
</tr>
<tr>
<td>IPP1</td>
<td>Overexpression</td>
<td>0.52</td>
<td>1.21</td>
<td>2.33</td>
<td>[36]</td>
</tr>
<tr>
<td>HDR1</td>
<td>Overexpression</td>
<td>0.75</td>
<td>2.5</td>
<td>3.33</td>
<td>[38]</td>
</tr>
<tr>
<td>HDR1</td>
<td>Down-regulation</td>
<td>0.05</td>
<td>0.09</td>
<td>1.8</td>
<td>[37]</td>
</tr>
<tr>
<td>GaFPPS</td>
<td>Overexpression</td>
<td>0.05</td>
<td>0.03</td>
<td>0.6</td>
<td>[42]</td>
</tr>
<tr>
<td>ADS</td>
<td>Overexpression</td>
<td>0.6</td>
<td>1.2</td>
<td>2</td>
<td>[45]</td>
</tr>
<tr>
<td>BRD2</td>
<td>Overexpression</td>
<td>0.94</td>
<td>2.14</td>
<td>2.26</td>
<td>[46]</td>
</tr>
<tr>
<td>ALDH1</td>
<td>Overexpression</td>
<td>0.8</td>
<td>22.6</td>
<td>2.83</td>
<td>[53]</td>
</tr>
<tr>
<td>HMGR, FPPS</td>
<td>Overexpression</td>
<td>0.8</td>
<td>25.6</td>
<td>3.2</td>
<td>[54]</td>
</tr>
<tr>
<td>HMGR, ADS</td>
<td>Overexpression</td>
<td>0.2</td>
<td>1.73</td>
<td>8.65</td>
<td>[47]</td>
</tr>
<tr>
<td>ADS-FPPS</td>
<td>Overexpression</td>
<td>0.91</td>
<td>2.44</td>
<td>2.68</td>
<td>[36]</td>
</tr>
<tr>
<td>CYP71AV1, CPR</td>
<td>Overexpression</td>
<td>0.71</td>
<td>0.98</td>
<td>1.38</td>
<td>[49]</td>
</tr>
<tr>
<td>FPPS, CYP71AV1, CPR</td>
<td>Overexpression</td>
<td>0.83</td>
<td>2.98</td>
<td>3.6</td>
<td>[50]</td>
</tr>
</tbody>
</table>
ADS, CYP71AV1, CPR
ADS, CYP71AV1, CPR, ALDH1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Competitive pathway key genes</th>
<th>Overexpression (mg/g FW)</th>
<th>Overexpression (mg/g FW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQS</td>
<td>Down-regulation</td>
<td>10</td>
<td>31.4</td>
</tr>
<tr>
<td>CPS</td>
<td>Down-regulation</td>
<td>2.32</td>
<td>3.6</td>
</tr>
<tr>
<td>CPS</td>
<td>Down-regulation</td>
<td>6.26</td>
<td>11.08</td>
</tr>
<tr>
<td>BFS</td>
<td>Down-regulation</td>
<td>6.26</td>
<td>11.08</td>
</tr>
<tr>
<td>GAS</td>
<td>Down-regulation</td>
<td>6.26</td>
<td>12.71</td>
</tr>
<tr>
<td>SQS</td>
<td>Down-regulation</td>
<td>6.26</td>
<td>10.70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>(mg/g FW)</th>
<th>(mg/g FW)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>6.4</td>
<td>15.1</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>27</td>
<td>3.4</td>
</tr>
</tbody>
</table>

1 CrHMGR and CrDXR genes were from C. roseus. 2 GaFPPS gene was from Gossypium arboreum.

4.1.2. Suppression of Competitive Metabolic Pathways

FPP has different metabolic flows in plants. It can not only be converted into amorpha-4,11-diene and enter the artemisinin biosynthesis pathway, but can also be converted into terpenoids such as β-farnesene, β-caryophyllene, germacrene A, squalene and epicedrol, which enter the biosynthesis pathway of other secondary metabolites (Figure 1). Therefore, by suppressing the flow of FPP into other competing pathways, more FPP can enter the artemisinin biosynthesis pathway and increase the yield of artemisinin [60,61] (Table 2).

SQS (squalene synthase) was the first key enzyme that catalyzed the sterol biosynthesis pathway. Using the hairpin RNA-mediated RNAi method to inhibit the expression of the SQS gene, the content of sterol was reduced by 37–58%, and the content of artemisinin was significantly increased, reaching 31.4 mg/g DW—3.14 times that of the control [57].

CPS (β-caryophyllene synthase) catalyzed the conversion of FPP to β-caryophyllene [62]. When the antisense RNA approach was adopted to down-regulate the expression of the CPS gene, the expression levels of the genes related to the artemisinin biosynthesis pathway were significantly up-regulated in transgenic plants. The results showed that the content of β-caryophyllene decreased by 40–62.7%, while the content of artemisinin increased by 54.9% [58].

Through RNA interference, the key enzyme genes, CPS, BFS, GAS and SQS, involved in the competitive pathway of artemisinin biosynthesis were gene-silenced. The results showed that the expression levels of key genes related to the artemisinin biosynthesis pathway (ADS, CYP71AV1, DBR2, ALDH1) were up-regulated, and the contents of artemisinin and DHAA were significantly increased. Compared with non-transgenic plants, the content of artemisinin in anti-CPS, anti-BFS, anti-GAS and anti-SQS transgenic plants rose by 77%, 77%, 103% and 71%, respectively [59].

In the biosynthesis pathway of artemisinin, RED1 (dihydroartemisinic aldehyde reductase) competed with ALDH1 to bind dihydroartemisinic aldehyde so that dihydroartemisinic aldehyde was converted into dihydroartemisinic alcohol, which means that artemisinin cannot be produced (Figure 1). If the expression level of the RED1 gene was reduced, more dihydroartemisinic aldehyde can enter the artemisinin biosynthesis pathway [63]. Ranjbar et al. analyzed the content of artemisinin in eight Artemisia species and the changes in the expression levels of key genes in the biosynthesis pathway of artemisinin. The results indicated that the content of artemisinin in A. annua was the highest, followed by Artemisia absinthium. This may be because the expression levels of the ADS and DBR2 genes in A. annua were higher than those in the other seven Artemisia species, while in A. absinthium, the expression of the ALDH1 gene was down-regulated and the expression of the RED1 gene was up-regulated, which reduced the artemisinin content in A. absinthium [64].
4.1.3. Regulation of Transcription Factors Expression

In the secondary metabolism of plants, transcription factors (TFs) can regulate the expression of a series of genes in metabolic pathways, and overexpression or suppression of these TFs is able to effectively regulate the accumulation of plant secondary metabolites [60,61]. A variety of TFs have been identified to be related to the biosynthesis of artemisinin, including WRKY, bHLH, AP2/ERF, bZIP, MYB, etc. (Table 3).

WRKY TF Family

WRKY TFs are one of the largest TF families in plants, with a conserved WRKYGQK sequence at the N-terminus and a typical zinc finger protein structure at the C-terminus [65]. WRKY TFs participate in the regulation of plant growth, development, senescence and response to environmental stress [66]. It has been reported that AaWRKY1, AaWRKY4, AaWRKY9, AaWRKY17, AaGSW1 and AaGSW2 are involved in the regulation of artemisinin biosynthesis (Table 3).

AaWRKY1 was the first cloned TF of A. annua [67]. AaWRKY1 activates the expression of the ADS, CYP71AV1 and DBR2 genes, and AaWRKY1 takes part in the jasmionic acid (JA) signaling pathway [68]. Overexpression of the AaWRKY1 gene increased the content of artemisinin by 1.3–2.0 times [60]. In AaWRKY4-overexpressed plants, the expression levels of the ADS, CYP71AV1, DBR2 and ALDH1 genes were significantly improved, and the production of artemisinin was increased by 35–50% [69]. AaWRKY9 was specifically expressed in the GSTs of A. annua, which responded to both light and JA signals, and positively regulated the biosynthesis of artemisinin. AaWRKY9 can bind to the promoters of the AaDBR2 and AaGSW1 genes to up-regulate the expression of artemisinin biosynthesis pathway genes, and the artemisinin production was increased by 1.6–2.2 times in AaWRKY9-overexpressed plants [70]. AaWRKY17 was also a TF that positively regulated artemisinin biosynthesis. Moreover, AaWRKY17 can activate the expression of two defense marker genes, PR5 (Pathogenesis-Related 5) and NHL10 (NDR1/HIN1-LIKE 10), and improve the tolerance of A. annua to Pseudomonas syringae, so AaWRKY17 could be used in the transgenic breeding of A. annua to improve artemisinin content and resistance to pathogens [71].

AaGSW1 (GLANDULAR TRICHOME-SPECIFIC WRKY 1) is a GSTs specific WRKY TF, which is regulated by AaMYC2 in the JA signaling pathway and AabZIP1 in the ABA signaling pathway. Overexpression of AaGSW1 activated the transcript levels of the ADS, CYP71AV1, DBR2, ALDH1 and ORA genes, and improved the content of artemisinin by 55–100% [72]. A recent study reported that AaGSW1 can bind to the promoters of TF AaTCP15 (teosinte branched1/cilyoidea/proliferating 15) and AaORA (octadecanoid-derivative responsive AP2-domain protein), forming the AaGSW1-AaTCP15/AaORA module network that regulates artemisinin biosynthesis through JA and ABA signaling transduction [73]. When AaGSW2 was overexpressed in A. annua, the density of GSTs was significantly increased in the overexpressed plants, and the artemisinin content went up 2-fold compared with the control [74].

bHLH TF Family

bHLH TFs are a class of TFs containing the basic helix-loop-helix (bHLH) domain that recognizes the E-box motif (CANNTG) of the promoter region [75]. According to the genome data, the bHLH TFs in A. annua have nearly 200 members, which makes it one of the largest TF families in A. annua [33]. So far, it has been reported that AabHLH1, AaMYC2 and AabHLH12 positively regulate the biosynthesis of artemisinin, while AabHLH2 and AabHLH3 are the negative regulators in artemisinin accumulation (Table 3).

AabHLH1 can bind to the promoters of the ADS and CYP71AV1 genes and positively regulate artemisinin biosynthesis. In AabHLH1-overexpressed plants, the expression levels of the ADS, CYP71AV1, DBR2 and HMGR genes are significantly up-regulated [76],
and the artemisinin content is increased 1.3-fold [77]. Moreover, the expression of AabHLH1 is induced by JA and AabHLH1 interacted with all nine AaJAZ proteins, so AabHLH1 is also involved in the regulation of JA-induced artemisinin biosynthesis [77].

AaMYC2 belongs to the bHLH TF superfamily and is the core factor of the JA signaling pathway, which can bind to the G-box motif of the CYP71AV1 and DBR2 gene promoters. Overexpression of AaMYC2 remarkably improved the transcription levels of the CYP71AV1 and DBR2 genes, and the content of artemisinin increased by 23.55% compared with the control, while in the RNAi plants of AaMYC2, the artemisinin content decreased to 54.81% of the control [78].

As a low-temperature-inducible TF, AabHLH112 positively regulates artemisinin biosynthesis through AaERF1. Yeast one-hybrid results showed that AabHLH112 cannot bind to the promoters of artemisinin biosynthesis-related genes, but could bind to the promoter of the AaERF1 gene [79]. AaERF1 is a member of the AP2/ERF TF family and can up-regulate the expression of artemisinin biosynthesis genes [80]. Overexpression of AabHLH112 significantly activated the transcript of the AaERF1 gene, which further promoted the expression of artemisinin biosynthesis genes. In AabHLH112-overexpressed plants, the artemisinin content reached 14.35 mg/g DW, which increased by 70.42% compared to the control (8.42 mg/g DW) [79].

AabHLH2 and AabHLH3 are MYC-type bHLH TFs, which suppress the expression of the ADS and CYP71AV1 genes by antagonizing AaMYC2 and participate in the negative regulation of artemisinin biosynthesis. Compared with the control, the artemisinin accumulation decreased by 27–66% and 21–61% in the overexpressed plants of AabHLH2 and AabHLH3, respectively. In contrast, the artemisinin content was increased by 42–87% and 35–60% in AabHLH2 and AabHLH3 RNAi plants, respectively [81]. As negative regulators, AabHLH2 and AabHLH3 may be good targets for prompting artemisinin production in A. annua by gene editing.

Table 3. Transcription factors involved in artemisinin biosynthesis pathway.

<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
<th>Expression Mode</th>
<th>Fold Change of Artemisinin Content</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WRKY TF family</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AaWRKY1</td>
<td>AaWRKY1 activates the expression of ADS, CYP71AV1, DBR2 genes.</td>
<td>Overexpression</td>
<td>1.30–2.00</td>
<td>[68]</td>
</tr>
<tr>
<td>AaWRKY4</td>
<td>AaWRKY4 prompts the expression of ADS, CYP71AV1, DBR2, ALDH1 genes.</td>
<td>Overexpression</td>
<td>1.35–1.50</td>
<td>[69]</td>
</tr>
<tr>
<td>AaWRKY9</td>
<td>AaWRKY9 positively regulates the expression of AaDBR2 and AaGSW1 genes.</td>
<td>Overexpression</td>
<td>1.60–2.20</td>
<td>[70]</td>
</tr>
<tr>
<td></td>
<td>AaWRKY17 activates the transcription of ADS, PR5 and NHL10 genes and enhances the resistance of A. annua to the pathogen P. syringae.</td>
<td>Down-regulation</td>
<td>0.79–0.85</td>
<td>[71]</td>
</tr>
<tr>
<td>AaGSW1</td>
<td>AaGSW1 is a GST-specific TF, and positively regulates the expression of CYP71AV1 and AaORA genes.</td>
<td>Overexpression</td>
<td>1.55–2.00</td>
<td>[72]</td>
</tr>
<tr>
<td>AaGSW2</td>
<td>As a GST-specific TF, AaGSW2 positively regulates the initiation of GST.</td>
<td>Overexpression</td>
<td>2.00</td>
<td>[74]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Down-regulation</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td><strong>bHLH TF family</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AabHLH1</td>
<td>AabHLH1 activates transcription of ADS and CYP71AV1 genes and responds to JA induction.</td>
<td>Overexpression</td>
<td>1.30</td>
<td>[77]</td>
</tr>
<tr>
<td>AaMYC2</td>
<td>AaMYC2 prompts the expression of CYP71AV1 and DBR2 genes and responds to JA induction.</td>
<td>Overexpression</td>
<td>1.24</td>
<td>[78]</td>
</tr>
<tr>
<td>AabHLH112</td>
<td>AabHLH112 promotes artemisinin biosynthesis by activating TF AaERF1.</td>
<td>Overexpression</td>
<td>1.70</td>
<td>[79]</td>
</tr>
</tbody>
</table>
### AaHLH2
AabHLH2 and AabHLH3 suppress the expression of ADS and CYP71AV1 genes by antagonizing AaMYC2, and negatively regulate artemisinin biosynthesis.

- **Overexpression**
  - 0.34–0.73
- **Down-regulation**
  - 1.42–1.87

### AaHLH3
AaMYB15

- **Overexpression**
  - 0.39–0.79
- **Down-regulation**
  - 1.35–1.60

### AaMYB17

- **Overexpression**
  - 1.1–1.2
- **Down-regulation**
  - ~1.52

### AP2/ERF TF family

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AaERF1</td>
<td>AaERF1 and AaERF2 activate the transcription of ADS and CYP71AV1 genes and respond to JA signaling.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaERF2</td>
<td>AaORA is a trichome-specific TF, which promotes the expression of ADS, CYP71AV1, DBR2 and AaERF1 genes, and improves the resistance of A. annua to the pathogen B. cinerea.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaTAR1</td>
<td>AaTAR1 activates the transcription of ADS and CYP71AV1 genes and regulates trichome development.</td>
<td>Overexpression</td>
</tr>
</tbody>
</table>

### bZIP TF family

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AaZIP1</td>
<td>AaZIP1 activates the transcription of ADS and CYP71AV1 genes through ABA signaling, indirectly promotes the expression of AaDBR2 and AaALDH1 genes through AaMYC2 and improves the drought resistance of A. annua.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaZIP9</td>
<td>AaZIP9 activates the expression of ADS gene and positively regulates artemisinin biosynthesis.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaHY5</td>
<td>AaHY5 positively regulates light-induced artemisinin biosynthesis through interacting with TF AaGSW1.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaTGA6</td>
<td>AaTGA6 takes part in artemisinin biosynthesis regulated by SA signaling and promotes artemisinin biosynthesis by activating the TF AaERF1.</td>
<td>Down-regulation</td>
</tr>
</tbody>
</table>

### MYB TF family

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AaMYB1</td>
<td>AaMYB1 participates in the biosynthesis of artemisinin and GA and regulates the trichome development.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaMIXTA1</td>
<td>AaMIXTA1 is a GST-specific TF, which regulates GST formation and cuticle biosynthesis.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaTAR2</td>
<td>AaTAR2 positively regulates GST development, as well as artemisinin and flavonoid biosynthesis.</td>
<td>Down-regulation</td>
</tr>
<tr>
<td>AaMYB16</td>
<td>AaMYB5 and AaMYB16 show the antagonism and regulate the development of GST by competitively binding to the AaHD1 promoter to form AaHD1-AaMYB5 or AaHD1-AaMYB16 complexes.</td>
<td>Down-regulation</td>
</tr>
<tr>
<td>AaMYB5</td>
<td>AaMYB17 is a GST-specific TF and positively regulates artemisinin biosynthesis.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaMYB17</td>
<td>As a GST-specific TF, AaMYB15 negatively regulates the biosynthesis of artemisinin by inhibiting the expression of AaORA gene.</td>
<td>Overexpression</td>
</tr>
<tr>
<td>AaMYB15</td>
<td>AaTLR1 and AaTLR2 interact with AaWOX1 to form the AaTLR1-AaWOX1-AaTLR2 complex to negatively regulate GST formation and artemisinin biosynthesis.</td>
<td>Overexpression</td>
</tr>
</tbody>
</table>

---

[80] [81] [82] [83] [84] [85] [86] [87] [88] [89] [90] [91] [92] [93] [94] [95]
<table>
<thead>
<tr>
<th>Other TF family</th>
<th>Description</th>
<th>TF Name</th>
<th>Gene Activation</th>
<th>Expression Fold</th>
</tr>
</thead>
<tbody>
<tr>
<td>AaZFP1</td>
<td>As a C2H2-type TF, AsZFP1 activates the transcription of <em>AalPP1</em> gene and prompts artemisinin biosynthesis.</td>
<td>AaZFP1</td>
<td>Overexpression</td>
<td>1.60 [96]</td>
</tr>
<tr>
<td>AaSPL9</td>
<td>As a C2H2-type TF, AsZFP1 activates the transcription of <em>AalPP1</em> gene and prompts artemisinin biosynthesis by activating TF AaHD1.</td>
<td>AaSPL9</td>
<td>Overexpression</td>
<td>1.33–1.60 [97]</td>
</tr>
<tr>
<td>AaSEP4</td>
<td>As a MADS TF, promotes artemisinin biosynthesis by directly activating TF AsGSW1.</td>
<td>AaSEP4</td>
<td>Overexpression</td>
<td>1.19–1.72 [98]</td>
</tr>
</tbody>
</table>

1 The change fold was calculated according to the data in the references.

**AP2/ERF TF Family**

The APELATA 2/ethylene response factor (AP2/ERF) TFs are one of the most important TF families in plants, and are involved in plant responses to biotic and abiotic stresses, as well as regulating various plant developmental and secondary metabolism processes [99]. AP2/ERF TF family members all contain the AP2 domain consisting of about 60 amino acid residues. According to the number of AP2 domains and whether they contain other domains, the AP2/ERF TF family is divided into AP2, ERF, DREB, RAV and Soloist 5 subfamilies [100,101].

AaERF1 and AaERF2 (Table 3), which respond to JA signaling, activate the transcription of the *ADS* and *CYP71AV1* genes by binding to the CBF2 and RAA motifs of these two gene promoters. Compared with the wild type, overexpression of AaERF1 prompted an increase in yield of artemisinin by 19–67% and overexpression of AaERF2 by 24–51% [80].

AaORA (Table 3) is a trichome specific AP2/ERF TF, which is expressed in GSTs and non-glandular T-shaped trichomes (TSTs). In AaORA-overexpressed plants, the expression levels of the *ADS*, *CYP71AV1*, *DBR2* and AaERF1 genes are all significantly up-regulated, and the contents of artemisinin and DHAA are increased by 40–53% and 22–35%, respectively [82].

In addition, overexpression of both AaERF1 and AaORA activated the transcription of defense marker genes, *PDF1.2* (PLANT DEFENSIN1.2) and *B-CHI* (BASIC CHITINASE), thereby enhancing the resistance of *A. annua* to the pathogenic fungus *Botrytis cinerea* [82,102]. A recent study reported that AaORA formed a transcriptional complex with the TF AaTCP14, which prompted the expression of *DBR2* and *ALDH1* genes, leading to an increased accumulation of artemisinin, and the AaORA-AaTCP14 complex was induced and activated by JA [103].

AaTAR1 (TRICHOME AND ARTEMISININ REGULATOR 1) (Table 3), as an AP2/ERF TF, directly binds to CBF2 and RAA motifs, activates the expression of the *ADS* and *CYP71AV1* genes, and participates in the positive regulation of glandular trichome development and artemisinin and cuticular wax biosynthesis. The contents of artemisinin and DHAA were enhanced by 22–38% and 69–130% in AaTAR1-overexpressed plants, respectively [83].

**bZIP TF Family**

bZIP (basic leucine zipper) TFs contain two specific motifs, one that is located at the C-terminal, which is composed of basic amino acids for binding to the target DNA sequence, and the other that is located at the N-terminal with the leucine zipper motif, which is required for the dimerization of bZIP TFs [104]. bZIP TFs participate in the regulation of plant growth and development to cope with various environmental stresses [105].

There are 75 bZIP family members in the *A. thaliana* genome, which are divided into 10 groups based on sequence similarity and conserved motifs [106]. The group A AtbZIP proteins contain ABRE (abscisic acid-responsive element) cis-elements, which are involved in ABA signal transduction [107]. ABA-dependent bZIP TFs can bind conserved
DNA sequences with ACGT core cis-elements, such as ABRE (ACGTG/TC), G-box (CACGTG), etc. [108]. AabZIP1, AabZIP9, AaHY5 and AaTGA6 are reported to be positive regulators of artemisinin biosynthesis (Table 3).

ABA treatment can prompt the artemisinin accumulation in A. annua [109]. AabZIP1 is a member of the group A bZIP family, and its expression is induced by abiotic stresses such as ABA, drought and high salt. AabZIP1 can bind to the ABRE cis-element of the ADS and CYP71AV1 gene promoters through the N-terminal CI domain and activate the expression of the ADS and CYP71AV1 genes. When AabZIP1 is overexpressed, the artemisinin content in transgenic plants showed a 1.5-fold increase [84].

A recent study reported that AabZIP1 can also directly bind to the promoter of the AaMYC2 gene and up-regulate the expression of AaMYC2, while AaMYC2 bound to the promoters of the AaDBR2 and AaALDH1 genes, which indicated that AabZIP1 indirectly activated the expression of the AaDBR2 and AaALDH1 genes through AaMYC2 and further improved artemisinin biosynthesis [85]. In addition, AabZIP1 can directly activate the transcription of cuticle wax biosynthesis genes AaCER1 and AaCYP86A1 and enhance the drought resistance of A. annua [85]. The results showed that the AabZIP1–AaMYC2 transcriptional module was a cross-talk between the ABA and JA signaling pathways in artemisinin biosynthesis, which provided a useful candidate gene for the genetic breeding of A. annua with a high artemisinin content and strong drought resistance.

As a member of the group C bZIP family in A. annua, AabZIP9 can bind to the “ACGT” cis-element of the ADS and CYP71AV1 gene promoters to up-regulate the expression of the ADS gene. Overexpression of AabZIP9 prompted the content of artemisinin and DHAA to increase by 23.2–67.1% and 34.5–92.8%, indicating that AabZIP9 was a positive regulator of artemisinin biosynthesis [86].

AaHY5, a member of the group H of bZIP TFs, regulated light-induced artemisinin biosynthesis by interacting with AaGSW1 in the WRKY family. When AaHY5 was overexpressed, the contents of artemisinin and DHAA were up-regulated approximately 2-fold [87].

SA (salicylic acid) treatment prompted the accumulation of artemisinin [110], and AaTGA6, a member of group D bZIP TFs, was involved in the artemisinin biosynthesis regulated by the SA signal [88]. As an important regulator of the SA signaling pathway, AaNPR1 up-regulated the expression of AaTGA6. AaTGA6 directly bound to the “TGACG” element of the AaERF1 gene promoter and activated the expression of key genes in artemisinin biosynthesis (ADS, CYP71AV1, DBR2, ALDH1). The artemisinin content was increased by 90–120% in AaTGA6-overexpressed plants, while AaTGA3 inhibited the binding of AaTGA6 to the AaERF1 gene promoter [88].

**MYB TF Family**

The MYB TFs contain a highly conserved DNA-binding domain: the MYB domain, which consists of about 52 amino acid sequence repeats (Rs), each of which form three α-helices [111]. According to the number of adjacent repeats, the MYB TFs can be divided into four classes: R2R3-MYB, 1R-MYB, 3R-MYB and 4R-MYB [111]. R2R3-MYB TFs mainly participate in cell morphogenesis [112], phytohormone signaling response [113], secondary metabolism regulation [114] and response to environmental stress [115]. The reported MYB TFs of A. annua all belong to the R2R3-MYB class. AaMYB1, AaMIXTA1, AaTAR2, AaMYB16 (AaMIXTA-like 2) and AaMYB17 positively regulate artemisinin biosynthesis, while AaMYB5, AaMYB15, AaTLR1 and AaTLR2 are the negative regulators in artemisinin accumulation (Table 3).

AaMYB1 promoted the glandular trichome initiation and artemisinin accumulation. Overexpression of AaMYB1 improved the density of glandular trichome, up-regulated the key genes of artemisinin biosynthesis, and significantly increased the content of artemisinin. Furthermore, both AaMYB1 and its orthologue, AtMYB61, participated in the regulation of trichome and root development, stomatal aperture and gibberellin biosynthesis [89]. AaMIXTA1 is predominantly expressed in the basal cells of GSTs and takes
part in GSTs formation and cuticle biosynthesis. Overexpression of AaMIXTA1 remarkably improved the accumulation of artemisinin [90]. AaTAR2 is also involved in the initiation and development of GSTs and regulated the secondary metabolism of terpenoids and flavonoids in *A. annua*. Overexpression of AaTAR2 significantly up-regulated the expression of key genes in artemisinin and flavonoid biosynthesis pathways, such as *HMGR, DXS, CYP71A1, DBR2* and *PAL, C4H, CHS, F3H*, and *DFR*. Furthermore, the HD-ZIP (homeodomain-leucine zipper) TFs, AaHD1 and AaHD8, prompted the transcription of AaTAR2 by binding to the L1 box (TAAAGATA) of the AaTAR2 promoter [91]. AaMYB17 was specifically expressed in GSTs of shoot tips. In AaMYB17-overexpressed *A. annua* plants, the density of GSTs was enhanced by 1.3–1.6 times, and the content of artemisinin was increased from 8 mg/g to 15 mg/g. In AaMYB17 RNAi plants, the GSTs density and artemisinin content were significantly reduced [93].

AaMYB15 is also a GSTs-specific TF, but unlike the previous R2R3-MYB TFs, AaMYB15 participates in the negative regulation of artemisinin accumulation. Overexpression of AaMYB15 resulted in the significant suppression of the expression of key genes for artemisinin biosynthesis, such as *ADS, CYP71AV1, DBR2* and *ALDH1*, and ultimately reduced the artemisinin content. The opposite results were observed in antisense AaMYB15 *A. annua* plants. Yeast one-hybrid results indicated that AaMYB15 could not bind to the promoters of these key genes, but rather directly bound to the promoters of *AaORA*. AaORA is an AP2/ERF TF that positively regulates artemisinin biosynthesis, and AaMYB15 inhibited artemisinin biosynthesis by suppressing the expression of the *AaORA* gene [94].

AaMYB5 and AaMYB16 are two antagonistic MYB TFs of *A. annua*, which regulate the development of GSTs. Overexpression of AaMYB5 inhibited the initiation of GSTs, resulting in the decline in artemisinin content, while AaMYB16 had the opposite effect; however, none of them could independently regulate the formation of GSTs, but played a regulatory role by competitively binding to the *AaHD1* gene promoter to form AaHD1–AaMYB5 or AaHD1–AaMYB16 complexes [92]. AaHD1 was an HD-ZIP TF that promoted the initiation of GSTs by directly activating the transcription of *AaGSW2*. In addition, JA was also associated with the AaHD1–AaMYB5/AaMYB16 regulatory network, which regulates the transcription of glandular trichome development-related genes through the AaJAZ8 protein [92].

A recent study reported two MYB TFs that negatively regulated GSTs development and artemisinin production: AaTLR1 and AaTLR2. Overexpression of AaTLR1 and AaTLR2 decreased the artemisinin content by 11.5–49.4% and 19–43%, respectively, while knockdown of the AaTLR1 and AaTLR2 genes resulted in the opposite effect. Yeast three-hybrid results showed that AaTLR1 and AaTLR2 interacted with AaWOX1 (WUSCHEL homeobox 1) to form the AaTLR1–AaWOX1–AaTLR2 complex, which negatively regulated the initiation of GSTs [95].

Other TF Families

In addition to the TFs discussed above, there are other TFs that also participate in the biosynthesis of artemisinin, such as AaZFP1, AaSPL9 and AaSEP4, all of which are positive regulators. AaZFP1 (zinc finger protein 1), as a C2H2-type TF, positively regulates the transcription of the AaIPPI1 gene by directly binding to the promoter of AaIPPI1. In transient AaZFP1 overexpressed *A. annua* plants, the artemisinin content was increased from 2.39 mg/g to 3.72 mg/g, which was 1.6 times the control [96]. AaSPL9 (SQUAMOSA promoter-binding protein like 9) can directly bind to the promoter of the *AaHD1* gene and activates the expression of the *AaHD1* gene. When AaSPL9 was overexpressed in *A. annua*, the density of glandular trichome increased by 45–60%, and the content of artemisinin went up by 33–60%—from 11.5 mg/g DW to 18.5 mg/g DW [97]. As a member of the MADS-box TF family, AaSEP4 is predominantly expressed in GSTs and activates the expression of TF AsGSW1 by directly binding to the CArG motif of the *AaGSW1* promoter. Overexpression
of AaSEP4 remarkably up-regulated the transcription levels of the AaGSW1, ADS, CYP71AV1, DBR2 and ALDH1 genes, and compared with the wild type, the content of artemisinin increased by 19–72% in AaSEP4-overexpressed plants [98].

Tables 2 and 3 summarize the current work of genetic engineering in A. annua to improve the generation of artemisinin. In general, the three strategies of “overexpression”, “suppression” and “global regulation” are adopted to achieve up-regulation of artemisinin accumulation. These strategies include overexpressing key enzymes that are upstream and downstream of the artemisinin biosynthesis pathway, blocking key genes in competitive metabolic pathways, and overexpressing or knocking-down TFs to globally regulate artemisinin biosynthesis. Based on the currently reported TFs of A. annua, the MYB TFs may act as key regulators, as they have the most binding sites in the promoters of artemisinin biosynthesis genes [116].

4.2. Genetic Engineering in Nicotiana Species

With Nicotiana as bioreactors, artemisinin had been successfully produced, but the yield of artemisinin was relatively low [21]. Tobacco had the advantages of large biomass and rapid growth, making it a suitable alternative plant for A. annua to study the heterologous biosynthesis of artemisinin [117] (Table 4).

When the ADS gene was expressed in Nicotiana tabacum, amorpha-4,11-diene, a precursor of artemisinin, can be produced in leaves at concentrations of 0.2–1.7 ng/g FW [118]. Through Agrobacterium-mediated transformation, five genes, including FPPS, ADS, CYP71AV1, DBR2 and ALDH1, were transformed into N. tabacum, and about 4 μg/g FW of amorpha-4,11-diene was accumulated in transgenic tobacco leaves [119]. Five genes (HMGR, ADS/mtADS, CYP71AV1, CPR and DBR2) derived from the MVA pathway and the artemisinin biosynthesis pathway were constructed into the same vector and overexpressed in N. tabacum. The highest concentration of artemisinin was 6.8 μg/g DW in transgenic plants [120].

Van Herpen overexpressed the HMGR, FPPS, ADS and CYP71AV1 genes of A. annua in Nicotiana benthamiana in order to produce the precursor of artemisinin—artemisinic acid in transgenic plants; however, artemisinic acid was not detected in the leaves, but the glycosylation product of artemisinic acid, artemisinic acid-12-β-diglucoside, was detected at a concentration of 39.5 μg/g FW [121]. This may be because when the key genes of artemisinin biosynthesis were heterologously expressed in N. benthamiana, glycosylated artemisinin precursors were mainly produced—with little free artemisinic acid or DHAA [122,123]. AaLTP3 (lipid transfer protein 3) and AaPDR2 (pleiotropic drug resistance 2) from A. annua prompted the accumulation of artemisinic acid and DHAA in leaf apoplasts of N. benthamiana and prevented the reflux of artemisinic acid and DHAA from the apoplast back into the cytoplasm [123].

Table 4. Comparison of artemisinin (or related products) biosynthesis in heterologous plants.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Plant Species</th>
<th>Expression System</th>
<th>Product Yield</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADS</td>
<td>N. tabacum</td>
<td>Nuclear</td>
<td>Amorpha-4,11-diene</td>
<td>[118]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.7 ng/g FW</td>
<td></td>
</tr>
<tr>
<td>FPPS, ADS, CYP71AV1, DBR2, ALDH1</td>
<td>N. tabacum</td>
<td>Nuclear</td>
<td>Amorpha-4,11-diene</td>
<td>[119]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 μg/g FW</td>
<td></td>
</tr>
<tr>
<td>HMGR, ADS/mtADS, CYP71AV1, CPR, DBR2</td>
<td>N. tabacum</td>
<td>Nuclear</td>
<td>Artemisinin</td>
<td>[120]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.8 μg/g DW</td>
<td></td>
</tr>
<tr>
<td>HMGR, FPPS, ADS, CYP71AV1</td>
<td>N. benthamiana</td>
<td>Nuclear</td>
<td>Artemisinic acid-12-β-glucoside</td>
<td>[121]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>39.5 μg/g FW</td>
<td></td>
</tr>
<tr>
<td>AACT, HMGS, HMGR, MVK, PMK, PMD, ID1, FPPS, ADS, CYP71AV1, CPR, aadA</td>
<td>N. tabacum</td>
<td>Chloroplast</td>
<td>Artemisinic acid</td>
<td>[124]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 μg/g FW</td>
<td></td>
</tr>
</tbody>
</table>
To address the issue of glycosylation in tobacco, researchers attempted to target key genes of the artemisinin biosynthesis to different cellular compartments, such as chloroplasts, mitochondria and the nucleus. A new transformation method, COSTREL (combinatorial super transformation of transplastomic recipient lines), was adopted to transform the complete pathway of artemisinic acid biosynthesis into chloroplasts of *N. tabacum*, and the content of the artemisinic acid in transgenic plants reached 120 μg/g FW [125]. In 2014, Kumar’s laboratory transformed 12 genes of the MVA pathway and artemisinic acid biosynthesis pathway into chloroplasts of *N. tabacum* through a biolistic approach, but the yield of artemisinic acid was low (100 μg/g FW) in transgenic tobaccos and the growth of transgenic tobaccos was stunted. These results suggested that it is necessary to disperse the genes of the artemisinin biosynthetic pathway to different cellular compartments [124]. In 2016, Kumar’s laboratory separately transformed these two biosynthesis pathways into chloroplast and nucleus genomes of *N. tabacum*, respectively, which did not interfere with tobacco growth. First, six genes (*AACT, HMGS, HMGR, MVK, PMK* and *PMD*) derived from the yeast MVA pathway were introduced into the chloroplast genome by a biolistic approach to increase the yield of IPP. Then, six key genes related to the artemisinin biosynthesis pathway (*IDI, FPPS, ADS, CYP71AV1, CPR* and *DBR2*) were transformed into the nuclear genome by the *Agrobacterium*-mediated method. Finally, DBR2, CPR and CYP71AV1 were transported into the chloroplast through the targeting of chloroplast transit peptides, and the artemisinin content in transgenic tobaccos reached 0.8 mg/g DW [126].

Although various methods have been used to improve the production of artemisinin in tobacco, due to the complexity of the heterologous expression and regulation of artemisinin biosynthesis genes, as well as the high level of glycosylation catalyzed by tobacco endogenous glycosyltransferase, the yield of artemisinin was still very low, but it is of great significance to study the heterologous biosynthesis of artemisinin in tobacco.

### 4.3. Genetic Engineering in Physcomitrella patens and Chrysanthemum morifolium

For more than two decades, *Physcomitrella patens* has become a model organism in plant biology, biotechnology and synthetic biology [130]. The *P. patens* genome was sequenced as early as 2008 [131], and the chromosome-level genomic data were released in 2018 [132]. In addition, *P. patens* had the advantages of a high homologous recombination rate, short growth cycle and large-scale culture, making it a green cell factory for metabolic engineering [133]. Different kinds of biopharmaceuticals, such as human complement factor H (FH), epidermal growth factor (EGF), hepatocyte growth factor (HGF), etc., have been successfully produced in *P. patens* [134]. The TXS (taxadiene synthase) gene from *T. brevifolia* was stably expressed in *P. patens*, and the content of taxadiene (the precursor of the anticancer drug paclitaxel) in the transgenic *P. patens* reached 0.05% FW [135].
In 2017, Ikram transformed five key genes of the artemisinin biosynthesis pathway (ADS, CYP71AV1, ADH1, DBR2 and ALDH1) into *P. patens* through homologous recombination. A high initial yield of 0.21 mg/g DW artemisinin was detected in transgenic *P. patens* after only three days of culture [127]. In 2019, Ikram heterologously expressed these five artemisinin biosynthesis genes in *P. patens* with different combinations, and the results showed that both combinations of three different genes, ADS–CYP71AV1–ADH1 and ADS–DBR2–ALDH1, can produce artemisinin, indicating that there may be endogenous enzymes in *P. patens* that can complement the biosynthesis pathway of artemisinin. In ADS–DBR2–ALDH1 transgenic lines, 0.04 mg/g DW of artemisinin was accumulated, and artemisinin B was detected at 1.74 μg/g FW in the liquid medium [128] (Table 4).

As a moss without vascular tissue, *P. patens* has less glycosylase compared with higher plants. Using *P. patens* as a chassis to produce artemisinin, it was less likely to endogenously modify the intermediate metabolites of artemisinin; however, the yield of artemisinin in *P. patens* was very low, so new molecular tools needed to be developed for *P. patens* to improve artemisinin accumulation.

Both *Chrysanthemum morifolium* Ramat and *A. annua* belong to the Compositae family and are characterized by a high content of sesquiterpenes and their precursors [136]. Firsov et al. transformed five artemisinin pathway genes, including HMGR, ADS, CYP71AV1, CPR and DBR2, into *Chrysanthemum* by Agrobacterium-mediated transformation; artemisinin was detected in transgenic lines by both GC-MS and TLC [129,137]. The results suggested that artemisinin biosynthesis genes can be expressed in transgenic *Chrysanthemum* to generate artemisinin (Table 4).

### 5. Challenges and Perspectives

Malaria still poses a threat to human health, and artemisinin is the most effective drug to treat malaria. The precursor of artemisinin, artemisinic acid, was generated in yeast through metabolic engineering, which produced the semi-synthesis of artemisinin and reached the level of industrial application. This was a huge advance in the research process of artemisinin biosynthesis, thereby transferring artemisinin production from plant to yeast. With inexpensive carbon sources as substrates, artemisinic acid was synthesized by fermentation of engineering yeast, but subsequent extraction and chemical synthesis are required before artemisinic acid can be converted into artemisinin.

To date, the main source of artemisinin is extracted from *A. annua* planted in fields, but the small amount of artemisinin in *A. annua* can hardly meet the needs of the pharmaceutical market. The heterologous expression of artemisinin biosynthesis genes in *Nicotiana* and *Physcomitrella* can stably generate artemisinin, but the yield of artemisinin is very low. Therefore, we should focus on *A. annua* itself to improve the yield of artemisinin through new biotechnologies, such as gene editing or molecular design breeding, etc.

The biosynthesis of artemisinin occurs from plant to yeast and then back to plant—e.g., *A. annua*, *Nicotiana*, *P. patens*, etc. In terms of artemisinin biosynthesis, it remains to be studied whether it is more advantageous to produce artemisinin by fermentation in microorganisms, or to extract artemisinin from *A. annua* cultivated in fields. Artemisinin is usually generated in the leaves of *A. annua* and *Nicotiana* by field planting. On the contrary, yeast and *P. patens* are commonly used for fermentation, and algae can even be fermented by light. Through high-cost fermentation, the desired product can be harvested under controlled conditions with shorter production cycles. On the other hand, the cost of planting in fields is lower, but the production cycle is longer. The following are several questions and challenges worth considering:

1. The final step of artemisinin biosynthesis was generally considered to be a photo-oxidation reaction. Can the artemisinic acid produced by the fermentation of engineered yeast be used for photochemical reactions in a large-scale photoreactor to increase the production of artemisinin?
2. The biosynthesis of artemisinin was mainly carried out in the GSTs on the surface of leaves. In the future, artemisinin will be produced in all tissues of *A. annua* leaves by new gene editing biotechnology and will not be limited to GSTs.

3. The molecular modification and directed evolution of key enzymes in artemisinin biosynthesis are performed by protein engineering methods to improve the activities of these key enzymes and improve the content of artemisinin.

4. *Chlamydomonas reinhardtii* is a model organism for the study of photosynthesis, known as “green yeast”. The genetic background of *C. reinhardtii* is clear, and its genome was released in 2007. Moreover, the chloroplast and nuclear transformation methods of *C. reinhardtii* have been established. At present, there is no report on the biosynthesis of artemisinin in *C. reinhardtii*. Is it possible to transform the genes of artemisinin biosynthesis into *C. reinhardtii* through codon optimization and generate artemisinin by photo-fermentation?

5. Various TFs have been reported to regulate artemisinin biosynthesis. Are these TFs acting independently or cooperatively, and which TF is the most important?

6. Conclusions

Based on the analysis of the *A. annua* genome, the biosynthesis pathway of artemisinin and its regulatory mechanism were elucidated, and the genes of key enzymes as well as TFs involved in artemisinin biosynthesis were isolated and identified. With microorganisms or plants as the chassis, the biosynthesis pathway of artemisinin was reconstructed to increase the content of artemisinin or its precursors in order to obtain artemisinin in an efficient and low-cost manner. Such research strategies provided a paradigm for the use of synthetic biology methods to produce natural products of medicinal plants. Many natural products have important physiological activities and are important sources for the development of new drugs. We are confident that by reconstructing and modulating the biosynthesis pathway through the synthetic biology approach, the efficient biosynthesis of rare and important natural products can be achieved, thereby addressing both the quality and quantity issues. In summary, there are broad application prospects for the promotion of high-value natural products from laboratory to industrialization.

**Author Contributions:** Conceptualization, L.Z., Y.Z. and H.J.; resources, Y.Z. and H.J.; data curation, L.Z., H.J. and Y.H.; writing—original draft preparation, L.Z.; writing—review and editing, M.W.; supervision, X.Z. and W.F.; funding acquisition, L.Z., X.Z. and W.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by National Key Research and Development Project (2017YFC1702800, 2019YFC1708802), Major Science and Technology Project in Henan Province (171100310500), Henan Province High-Level Personnel Special Support “ZhongYuan One Thousand People Plan” (ZYQR201810080), Training Program for Young Teachers in Colleges and Universities of Henan Province (2021GGJS086).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** We thank Kai Tan for assisting in the preparation of this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


Molecules 2022, 27, 6888


Molecules 2022, 27, 6888


Molecules 2022, 27, 6888


Molecules 2022, 27, 6888


Molecules 2022, 27, 6888


