Agrotechniques in Industrial Crops

Journal Homepage: https://atic.razi.ac.ir

Decoding the Biosynthetic Pathway of the Alkaloid Morphine with Bioinformatics

Nazila Bagheri[*1], Alireza Tarinejad[1], Karim Hasanpur[2], Mohammad Majidi[1]

[1]*Agricultural Biotechnology Department, Faculty of Agriculture, Azarbaijan Shahid Madani University, Tabriz, Iran*
[2]*Department of Animal Sciences, University of Tabriz, Tabriz, Iran*

ARTICLE INFO

Original paper

Article history:
Received: x Month 202x
Revised: x Month 202x
Accepted: x Month 202x

Keywords:
Analysis
Bioinformatics
Hub genes
Morphine
Poppy
Secondary metabolites

ABSTRACT

In 1803, the opium poppy was the source of Morphine, the first alkaloid extracted. Due to its diverse use and therapeutic applications, it is considered the most notable alkaloid and accounts for 42 out of all alkaloid substances. This study aimed to use bioinformatics techniques to investigate the biosynthesis pathway of morphine. It included the compilation of nine genes associated with this pathway, based on a thorough literature review. The genes were later confirmed with the NCBI BLAST tool. For examining gene interactions, the research used STRING, and Cytoscape was utilized to visualize the molecular interaction network. Additionally, CytoHubba was applied to pinpoint hub proteins in this network. The hub genes were examined for enrichment with the Kyoto Encyclopedia of Genes and Genomes (KEGG) using STRING, while Gene Ontology (GO) analysis was conducted through gprofiler. Furthermore, the promoter regions of important genes were analyzed using MEME. The metabolic processes involved in morphine production highlight that the gene network associated with the morphine pathway has wider functions beyond merely generating primary metabolites. An examination of the KEGG pathway highlighted the importance of metabolic pathways and the production of secondary metabolites. Additionally, a review of the promoter suggested that signal transduction might be involved in morphine synthesis. The main genes involved in the production of morphine are linked to several key plant pathways, given that morphine is categorized as a secondary metabolite. This study employs various bioinformatics tools to pinpoint and evaluate gene interactions and metabolic pathways, providing a better understanding of how morphine alkaloids are synthesized. This approach could help develop new methods for producing and extracting morphine, as well as improve agricultural practices related to medicinal plants.

DOI: 10.22126/ATIC.2025.11523.1182

1. Introduction

The poppy (*Papaver somniferum* L.), belonging to the Papaveraceae family, is globally acknowledged as an important medicinal plant. It is especially known for its secondary benzylisoquinoline alkaloids (BIAs), such as morphine, codeine, thebaine, and oripavine. These substances offer a range of therapeutic benefits, including pain management, addiction recovery, and cancer treatment (Aghaali and Naghavi, 2024). In reality, BIAs can help alleviate pain, address addiction issues, and support cancer treatment; however, improper use can result in addiction (van Dorp *et al.*, 2007). Various types of poppies possess distinct profiles of biologically active compounds, with certain varieties containing greater levels of morphine and codeine, while others are more abundant in thebaine and oripavine (Desgagné-Penix *et al.*, 2012). Morphine is a powerful pain reliever and anesthetic, but it has considerable side effects and a potential for addiction. Its structure includes a long pentacyclic core with five linked chiral centers, one of which is a quaternary benzylic carbon, creating the morphinan framework. Because of its distinctive characteristics, morphine is seen as an intriguing substance for assessing contemporary synthetic methods (Tolchin *et al.*, 2023). The production of secondary metabolites in plants is a complicated process that is controlled by various genes and affected by many different factors (Tahmasebi *et*

* Corresponding author.
E-mail address: bagheri12255@gmail.com

al., 2019). Secondary metabolites help plants react to various environmental changes (Fazili *et al.*, 2022). Manufacturers aim to investigate the molecular processes involved in the production of secondary metabolites and utilize this understanding to develop pharmaceuticals (Malik *et al.*, 2023). The development of BIAs is a complex process influenced by genetic factors and various other elements, emphasizing the need to understand the underlying molecular mechanisms to improve production (Garg *et al.*, 2015). As the number of plant genome sequences and post-genomic research increases, a significant challenge is to utilize this genomic data to improve our understanding of the molecular mechanisms that contribute to the development of complex traits. A gene network depicts the relationships between genes, with nodes representing the genes and edges indicating their interactions, underscoring the importance of network analysis in plant biology. Genes that interact, either directly or indirectly, are likely to play similar biological roles. These networks uncover potential links among genes and enable a systematic investigation of the molecular mechanisms that underlie biological processes. By identifying and assessing the significance of nodes within the network, we can pinpoint and emphasize the key components of biological networks (Liu *et al.*, 2020).

Cluster analysis is a technique employed in networks to identify functional modules and predict protein complexes and biomarkers within the network (Li *et al.*, 2017). The synthesis of these alkaloids depends on a complex network of genes and enzymes distributed among different cell types in the plant. Key genes associated with BIA biosynthesis have been identified through various genomic and transcriptomic techniques. Furthermore, several enzymes that play a role in the sanguinarine biosynthesis pathway were found to be among the most abundant transcripts in cell cultures exposed to elicitors (Aykanat and Türktaş, 2024). A research study utilized next-generation sequencing to identify genes associated with the biosynthetic pathway of secondary metabolites in the Abu Jahl variety of watermelon. RNA was extracted from the fruit tissues of this watermelon variety, and sequencing was performed on the Illumina Hiseq 2500 platform. The bioinformatics analysis incorporated a unique integration with the Evidential-gen software and functional analysis through the KAAS database. Out of 21,952,885 high-quality sequences, 55,311 individual genes were identified and submitted to various databases. The KAAS database indicated that 17,359 of these genes were annotated across 134 different plant pathways. Notably, within the pathways related to significant secondary metabolites in the fruit tissue, 39 monogenes and 8 ortholog genes were associated with the triterpene and sesquiterpene pathways. The goal of this transcriptome analysis of the medicinal plant is to pinpoint the genes involved in the biosynthetic pathways of secondary metabolites. This research opens up opportunities for various applications, including the potential engineering of biosynthetic pathways for herbal medicinal compounds (Dorafshan *et al.*, 2019).

In a particular research study, researchers produced 700 Mb of transcriptome sequences using the Pac Bio platform, known for its effective use of the natural DNA replication process. After filtering out low-quality data and removing technology-related adapters, they successfully assembled 302 Mb of high-quality long-read sequences into 120,926 contigs, with an average length of 3,117 bases (Manni *et al.*, 2021). Most of the sequences were found to be between 3000 and 3500 base pairs long. To investigate the genes associated with secondary metabolite biosynthesis in the medicinal poppy, heat maps were generated by analyzing orthologous genes. A phylogenetic tree was also created to examine the evolutionary connections among these genes, and the alkaloid levels were assessed. Furthermore, 137 transcripts related to secondary metabolite biosynthesis, which showed different expression levels at various growth stages of the plant, were analyzed as well (Filiault *et al.*, 2018). The advent of plant genome sequencing has rendered network analysis an effective approach for understanding gene interactions and their roles in basic traits. Specifically, for opium poppy, network analysis has yielded significant information regarding the pathways that contribute to morphine production. This research seeks to identify novel genes and networks associated with the morphine biosynthetic pathway by leveraging different databases and software tools. The findings from this study will improve our understanding of the relationships between biosynthetic genes and their functional pathways, potentially benefiting genetic engineering in both research and industrial contexts.

2. Materials and methods

This experiment was conducted in 2023 at the Faculty of Agriculture, Azerbaijan Shahid Madani University.

2.1. Morphine pathway genes and protein data

Information on the quantity and names of genes related to the morphine biosynthetic pathway in *P. somniferum* was collected. A total of nine genes linked to morphine production were identified, and their protein sequences were retrieved from NCBI. These protein sequences were then examined for protein-protein interactions (PPI) using the STRING database. Details about the nine genes involved in the morphine biosynthetic pathway, including their gene IDs and descriptions, are available in Table 1.

2.2. Protein-protein interaction (PPI) network analysis

Nine protein groups related to morphine biosynthesis were entered into the STRING program (version 10) (https://stringdb.org) to forecast their functional interactions with other proteins in *P. somniferum*. The STRING database contains both established and anticipated protein-protein interactions (PPIs) for more than 2000 species. The Cytoscape program (version 3.9.1) was subsequently utilized to visualize the PPI network. The CytoHubba plugin (version 0.1) within Cytoscape was used to assess the central proteins among all nodes. Four hub ranking methods from CytoHubba—local-based degree (Deg), maximum clique centrality (MCC), maximum neighbor component (MNC), and maximum neighbor component density (DMNC)—were analyzed to evaluate the scores of the nodes in the network, considering the connections between nodes and their immediate environments. Important features within the PPI network were identified using the CytoHubba plugin to highlight significant hubs with high relevance in the network's structure (Ghorbani *et al.*, 2023). The identified hub genes were then input into STRING software to forecast a subnet, which was subsequently visualized using Cytoscape software.

2.3. Gene ontology analysis and pathway enrichment in the subnetwork

To pinpoint important biochemical pathways associated with the hub genes and their subnetworks, the accession numbers of the genes related to these subnetworks were entered into the gprofiler database (https://biit.cs.ut.ee/gprofiler/gost). The following functional analysis and enrichment section, which came after the network visualization, played a vital role in this study. The results were related to molecular functions (MF) and cellular components (Desgagné-Penix *et al.*, 2012), Biological processes (BP) were gathered for the classification of gene ontology (GO) (Karimizadeh *et al.*, 2019). Furthermore, KEGG pathway data was acquired through STRING for enrichment analysis, with a statistical significance threshold established at $p<0.05$.

2.4. Hub gene promoter motif analysis

The 1 kb regions located upstream of the hub gene were obtained from the Ensembl plant web service (https://plants.ensembl.org). Conserved motifs within these sequences were detected using MEME Suite version 5.4.1 (meme.nbcr.net/meme/intro.html) with the default settings, although the P and E thresholds were adjusted to be below 0.01. Furthermore, known cis-regulatory elements (CRE) were identified using the JASPAR CORE2022 database, also with default settings but with P and E thresholds set to under 0.0001 (Vorontsov *et al.*, 2024). The locations of important motifs were subsequently analyzed.

2.5. Investigation of the most important genes and key enzymes in the morphine biosynthesis pathway

To reach this objective, KEGG mapper color (https://www.genome.jp/kegg/mapper/color.html) was employed. A diagram depicting the biosynthetic process was generated, and the important genes and enzymes related to this pathway were analyzed. Two morphine synthesis pathways (the main pathway and the secondary pathway) have been recognized.

3. Results and discussion

3.1. PPI network analysis

The purpose of this research was to discover new genes and networks linked to the morphine biosynthetic pathway in *P. somniferum*, as well as to improve the understanding of the relationship between biosynthetic genes and functional pathways. The STRING database generated 121 nodes and 672 edges (Fig. 1) based on the overall interactions. These nodes and edges were then imported into Cytoscape to create a protein-protein interaction (PPI) network. The

CytoHubba analysis identified 9 hub genes with the most interactions (Table 1). A subnetwork was formed from these hub genes, and all genes within this subnetwork are provided along with their annotations and IDs (Fig. 2). According to the annotations, LOC113321914 is an uncharacterized protein. Two of the key genes, LOC113322169 and LOC113336808, belong to the Bet_v_1 domain-containing protein family. This family belongs to the PR-10 protein group, which is associated with pathogenicity and exhibits structural resemblances to various protein superfamilies (Radauer *et al.*, 2008). Bet v 1 has a prominent hydrophobic cavity that extends through the middle of the protein, playing a vital role in its capacity to bind to ligands. Importantly, Bet v 1 shows ribonuclease activity, evidenced by its capability to break down plant RNA (Grutsch *et al.*, 2017). Additionally, various isoforms of Bet v 1 display different degrees of structural flexibility and vulnerability to proteolytic degradation, potentially affecting their allergenic characteristics. Research indicates that this protein binds to several key physiological ligands, such as fatty acids, flavonoids, and cytokinins, with a binding affinity in the low micromolar range (Mogensen *et al.*, 2002). As a result, the Bet v 1 domain-containing protein serves as a versatile structure for binding large ligands and may function as a transporter for plant steroids. Two key genes (LOC113347789 and LOC113311621) are associated with the Fe2OG dioxygenase domain-containing protein and are part of the iron/ascorbate-dependent oxidoreductase family (Aglas *et al.*, 2020).

Table 1. Ranking of hub genes identified in morphine biosynthesis using CytoHubba computational algorithm, including MNC, MNCC, Degree and MCC

NO	String name	Gene ID	Query ID	Selected by	Annotation
1	A0A4Y7LGY7	LOC113322169	C5167_047537	MCC, MNC, Degree	Bet_v_1 domain-containing protein
2	A0A4Y7ILV1	LOC113347789	C5167_018133	DMNC	Fe2OG dioxygenase domain-containing protein; Belongs to the iron/ascorbate-dependent oxidoreductase family
3	A0A4Y7IDB6	LOC113311621	C5167_039857	MNC, MCC, Degree	Fe2OG dioxygenase domain-containing protein; Belongs to the iron/ascorbate-dependent oxidoreductase family
4	B6E2Y6, SAT	LOC113321357	c5167_047541	DMNC	Salutaridinol 7-O-acetyltransferase
5	A0A4Y7LKW4	LOC113322262	C5167_047536	MCC, MNC, Degree	Salutaridine reductase
6	A0A4Y7IEY3	LOC113336808	C5167_040371	MCC	Bet_v_1 domain-containing protein
7	A0A4Y7LKP9	LOC113321914	C5167_047533	DMNC	Uncharacterized protein
8	A0A4Y7K772	LOC113293958	C5167_031316	MCC, DMNC	Aldo_ket_red domain-containing protein
9	A0A4Y7LIF6	LOC113322712	C5167_047531	MNC, Degree	Aldo_ket_red domain-containing protein

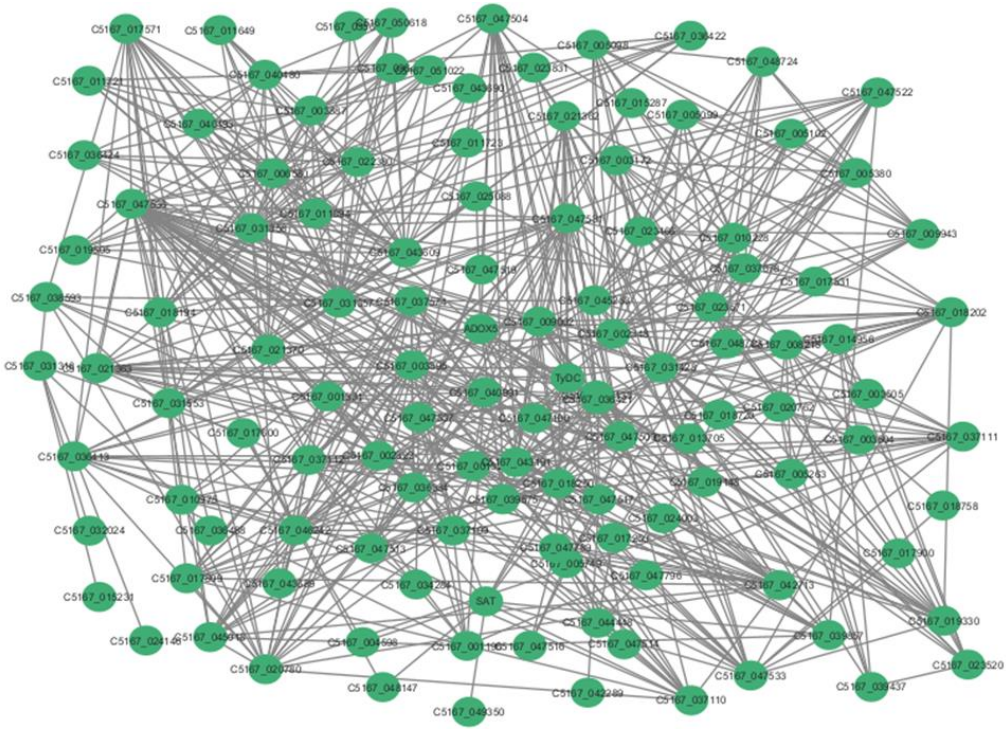


Figure 1. PPI Network of the morphine biosynthesis genes in opium poppy using Cytoscape software and STRING database (Genes related to morphine biosynthesis were first placed in the String database and then transferred to Cytoscape software to examine the relationships of these genes).

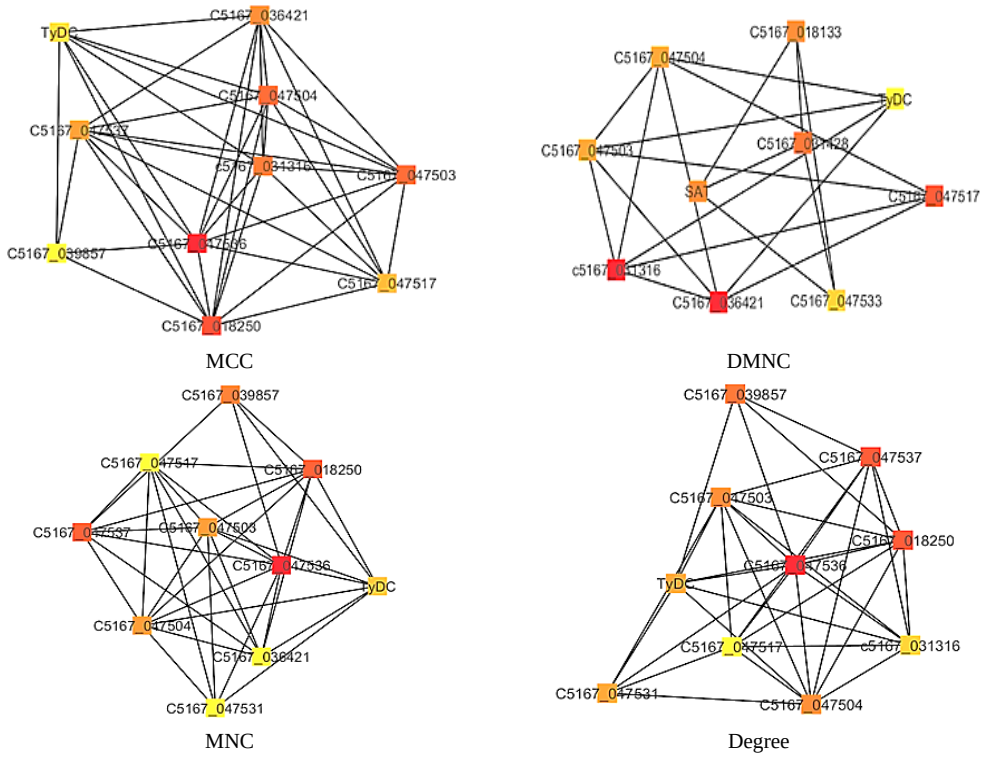


Figure 2. Protein-protein interaction [PPI] network of hub genes with the four methods (MCC, DMNC, MNC, and Degree)

Fe (II)/2-oxoglutarate-dependent dioxygenases (Fe2OG) are a varied group of enzymes that facilitate different oxidation reactions by utilizing iron as a cofactor and 2-oxoglutarate as a substrate. These enzymes are crucial for numerous biological processes, such as collagen synthesis, fatty acid metabolism, sensing low oxygen conditions, and the dimethylation of DNA and histones (Kuiper and Vissers, 2014). Proteins that contain the Fe2OG dioxygenase domain are part of the iron/ascorbate-dependent oxidoreductase family, which necessitates Fe (II) and frequently ascorbate as cofactors for their enzymatic functions. While many Fe2OG dioxygenases share a similar catalytic mechanism, they differ in their substrate specificities and roles. For instance, the Ofd2 protein from Schizosaccharomyces pombe, which belongs to the AlkB subfamily, exhibits decarboxylase activity that is enhanced by histones, especially H2A (Osorio-Concepción *et al.*, 2021).

Certain Fe2OG dioxygenases, especially those that regulate hypoxia-inducible factor (HIF), are crucial in tumor biology and could serve as promising targets for cancer treatment. In conclusion, proteins containing the Fe2OG dioxygenase domain are multifunctional enzymes that participate in a wide range of biological activities. Their varied functions and substrate preferences render them significant targets for research in fields like epigenetics, DNA repair, and cancer biology. The dependence of these enzymes on iron and ascorbate underscores the necessity of sustaining adequate iron levels and ascorbate availability for cellular metabolism and signaling (Kuiper and Vissers, 2014). Two more hub genes, LOC113293958 and LOC113322712, belong to the Aldo_ket_red domain-containing protein family. Proteins that contain the aldo keto reductase (AKR) domain are members of a vast superfamily of NAD(P)(H)-dependent oxidoreductases present in a range of organisms (Huacachino *et al.*, 2024). These enzymes facilitate redox reactions that play a crucial role in biosynthesis, intermediary metabolism, and detoxification. They exhibit a wide range of substrate specificity, including glucose, steroids, glycosylation products, lipid peroxidation products, and different environmental pollutants (Shanbhag and Bhowmik, 2024).

Proteins that have the AKR domain are crucial for many biological functions, including detoxifying harmful substances and participating in biosynthesis and metabolism. Their significance is highlighted by their occurrence in a variety of organisms, from bacteria to plants and animals, and their role in different cellular processes. The evolutionary variety and functional capabilities of the AKR superfamily make it an important focus for research aimed at understanding

cellular regulation, protective mechanisms, and possible uses in industrial catalysis (Krishnamurthy *et al.*, 2022).

A significant gene (LOC113321357) belongs to the salutaridinol 7-O-acetyltransferase family. Salutaridinol 7-O-acetyltransferase (SalAT) is crucial for the production of morphine in *P. somniferum*. It facilitates the transformation of salutaridinol into salutaridinol-7-O-acetate, which is the direct precursor to morphine (Yucebilgili Kurtoglu and Unver, 2021). SalAT is associated with salutaridine reductase (SalR), suggesting the creation of an enzyme complex within the morphine biosynthesis process. This relationship could clarify why salutaridine levels rise when SalAT is suppressed via RNA interference. Altering the SalAT gene, either by boosting or suppressing its activity, impacts the alkaloid concentrations in opium poppy; boosting the gene leads to elevated levels of morphine, codeine, and thebaine, whereas suppression results in increased salutaridine (Allen *et al.*, 2008). As a result, SalAT is essential for the synthesis of morphinan alkaloids and has been the subject of extensive research aimed at boosting alkaloid yields in *Papaver* species. Its significance is underscored by its application in metabolic engineering initiatives aimed at enhancing alkaloid production in *P. somniferum* and *P. bracteatum*. The activity of this enzyme is influenced by feedback inhibition from thebaine, which could serve as a potential target for increasing opioid alkaloid production in genetically altered yeast strains (Ozber *et al.*, 2023).

Gene LOC113322262 belongs to the salutaridine reductase family. Salutaridine reductase (SalR) plays a vital role in the production of morphine and codeine in the opium poppy. It facilitates the reduction of the C-7 keto group in salutaridine, leading to the creation of salutaridinol, which is a key intermediate in the synthesis of morphine (Higashi *et al.*, 2011). A study of the genes and pathways related to morphine production in *P. somniferum* has revealed a group of essential genes with diverse functions. These findings underscore the significant molecular roles and pathways that could significantly influence the morphine biosynthesis process in the opium poppy. Recognizing these important genes that regulate various biological processes highlights their potential significance in navigating the intricate interactions within the morphine biosynthesis pathway.

3.2. Gene ontology exploration and pathway enrichment for subnetwork genes in P. somniferum *associated with morphine production*

Gene Ontology (GO) analysis is a well-established method used to identify genes and their associated products, as well as to highlight biological characteristics obtained from large-scale genomic or transcriptomic data. This includes aspects of molecular function (MF) and cellular component (Desgagné-Penix *et al.*, 2012), and biological process (BP) (Jha *et al.*, 2024). The results from the GO analysis and pathway enrichment are summarized here. The main GO terms associated with biological processes, which represented 50% of the identified genes, showed significant enrichment in areas like metabolic processes, cellular processes, organic matter metabolism, and cellular metabolism. Our results indicate that metabolic processes play a vital role in morphine production. The biological functions, which involve specific interactions with ligands or the structures of gene products, should be considered the molecular functions of a gene product. The biochemical function of a gene product is defined as a molecular function that encompasses specific interactions with ligands and their structures. Furthermore, this definition also includes the potential capabilities of an individual gene product or a combination of gene products (Aleksander *et al.*, 2023). A detailed analysis shows that the gene network responsible for the morphine pathway affects processes that go beyond its known metabolic functions related to secondary metabolite production. Although the main role of the morphine pathway is established, the sub-hub gene network appears to serve as a multifunctional element engaged in various biological activities. This indicates that the sub-network plays a role not only in the primary metabolic pathway but also in other processes, emphasizing its extensive functional capabilities. These results reveal the intricate and adaptable nature of the gene network involved in morphine biosynthesis and offer important insights into potential regulatory mechanisms and broader biological consequences.

In the classification of cellular components, important Gene Ontology (GO) terms, which represent over 55% (≥55%) of the identified genes, highlight a strong link to anatomical structures. This indicates a specific relationship between the genes in the

subnetwork and the cellular components essential for morphine production and its regulation. The notable presence of anatomical units implies that these genes actively shape the cellular environment and are likely involved in creating specialized compartments necessary for morphine synthesis. From a molecular function perspective, the primary Gene Ontology (GO) terms, which make up more than 50% (≥50%) of the identified genes, show significant enrichment in activities related to catalysis, ion binding, oxidoreductase functions, and transferase functions. These molecular functions are linked to important steps in the morphine production pathway. The presence of catalytic activity suggests that an enzyme is crucial for the major biochemical changes necessary for morphine synthesis. The activities associated with ion binding and oxidoreductase indicate that the genes in this subnetwork probably contribute to redox reactions or ion interactions that are essential for morphine biosynthesis. Additionally, transferase activity is involved in moving functional groups between different molecules. These components fall under the category of molecular functions. Biological processes, including methylation and the metabolic pathways related to alkaloid synthesis, are classified as biological processes (Table 2). In addition to Gene Ontology (GO) analysis, KEGG pathway enrichment analysis provides further insight into the relationships between this subnetwork and morphine biosynthesis in *P. somniferum* (Fig. 3). This evaluation highlights specific pathways that are significantly affected by the genes within the identified subnetwork, providing important insights into the functional structure related to morphine production. The KEGG pathway enrichment analysis indicates that the main pathways, which encompass more than 45% of the identified genes, focus on metabolic pathways and the synthesis of secondary metabolites. This finding supports the important and well-established function of the morphine biosynthetic pathway. Metabolic pathways, which are a key enriched category, emphasize the significant involvement of subnetwork genes in various biochemical processes essential for morphine synthesis. The production of secondary metabolites is particularly associated with morphine generation, as it represents a vital segment of the metabolic network responsible for producing specialized compounds (Bharadwaj *et al.*, 2021). The enhanced pathways effectively demonstrate the important roles of the subnetwork genes in supporting the morphine biosynthesis pathway. The focus on metabolic pathways and the production of secondary metabolites underscores the vital contributions of the identified genes in managing the complex series of reactions that lead to morphine synthesis. This integrated examination of GO terms and KEGG pathways provides an overview of the functional aspects of the subnetwork and confirms its critical involvement in the intricate mechanisms that control morphine biosynthesis in opium poppy.

Table 2. Gene ontology exploration and pathway enrichment for subnetwork genes associated with morphine production

ID	Source	Term ID	Term name
1	GO:MF	GO:0003824	Catalytic activity
2	GO:MF	GO:0008483	Transaminase activity
3	GO:MF	GO:0008168	Methyltransferase activity
4	GO:MF	GO:0070279	Vitamin B6 binding
5	GO:MF	GO:0004185	Serine-type carboxypeptidase activity
6	GO:MF	GO:0005506	Iron ion binding
7	GO:BP	GO:0009820	Alkaloid metabolic process
8	GO:BP	GO:0032259	Methylation
9	GO:MF	GO:0008757	S-adenosylmethionine-dependent methyltransferase activity
10	GO:MF	GO:0016769	Transferase activity, transferring nitrogenous groups
11	GO:MF	GO:0016741	Transferase activity, transferring one-carbon groups
12	GO:MF	GO:0016491	Oxidoreductase activity
13	GO:MF	GO:0016705	Oxidoreductase activity, acting on paired donors, with incorportion or reduction of molecular oxygen
14	GO:MF	GO:0008171	O-methyltransferase activity
15	GO:MF	GO:0004497	Monooxygenase activity
16	GO:MF	GO:0016706	2-oxoglutarate-dependent dioxygenase activity
17	GO:MF	GO:0016717	Oxidoreductase activity, acting on paired donors, with oxidation of a pair of donors resulting in the reduction of molecular oxygen to two molecules of water
18	GO:MF	GO:0019842	Vitamin binding
19	GO:MF	GO:0030170	Pyridoxal phosphate binding
20	GO:MF	GO:0030782	(S)-tetrahydroprotoberberine N-methyltransferase activity
21	GO:MF	GO:0070008	Serine-type exopeptidase activity
22	GO:MF	GO:0102802	Thebaine 6-O-demethylase activity
23	GO:BP	GO:0019748	Secondary metabolic process
24	GO:BP	GO:0009821	Alkaloid biosynthetic process
25	GO:BP	GO:0033075	Isoquinoline alkaloid biosynthetic process
26	GO:BP	GO:1901376	Organic heteropentacyclic compound metabolic process
27	GO:BP	GO:1901378	Organic heteropentacyclic compound biosynthetic process
28	GO:BP	GO:0097295	Morphine biosynthetic process
29	GO:BP	GO:0071272	Morphine metabolic process
30	GO:BP	GO:1901564	Organonitrogen compound metabolic process

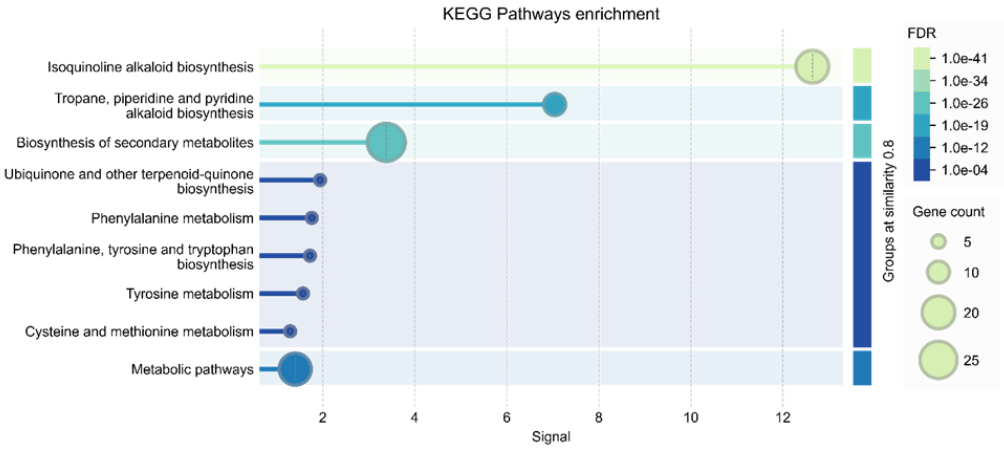
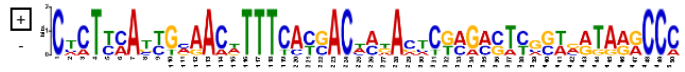
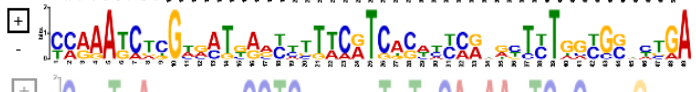


Figure 3. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis on the hub genes in morphine biosynthesis using STRING version10. (In this analysis, Isoquinoline, Tropen, piperidine and pyridine alkaloid biosynthesis, Biosynthesis of secondary metabolites, Ubiquinone and another terpenoid-quinone biosynthesis, phenylalanine metabolism, Phenylalanine, tyrosine and tryptophan biosynthesis, Tyrosine metabolism, Cysteine and methionine metabolism, Metabolic pathways play a role).

3.3. Hub gene promoter motif analysis

The conserved motifs and shared cis-regulatory elements (CREs) of important genes were examined using UFRs identified in Ensemble Plants. MEME revealed eight primary functions within the promoter gene. Details about their E-value, width, sites, and logos are presented in Table 3. The E-value of a motif is determined based on its log-likelihood ratio, width, number of sites, background letter frequencies, and the size of the training dataset. This E-value estimates the expected number of motifs with the given log-likelihood ratio (or higher), width, and site count that could be found in a similarly sized set of random sequences. The term "sites" refers to the total number of positions that contributed to the formation of the motif, while the width of the motif indicates the size of a pattern with a fixed width, as MEME motifs do not allow for any gaps. Subsequently, the placement of the motif is illustrated in (Fig. 4). This illustration shows the positions of motif sites, with each block representing the site's location and strength. The height of each block indicates the site's significance, with taller blocks denoting greater importance. This height is calculated using the negative logarithm of the site's p-value, capped at a maximum height that corresponds to a p-value of 1e-10. For complementary alphabets such as DNA, sites on the positive strand are represented above the line, while those on the negative strand are shown below. The position p-value is defined as the likelihood that a random sequence, of the same length and following the background distribution, would yield a motif score equal to or greater than the highest score observed in the analyzed sequence. Notably, the highest p-value is linked to gene c5167-018133, which is associated with the Fe2OG dioxygenase domain-containing protein from the iron/ascorbate-dependent oxidoreductase family; in contrast, the lowest p-value is related to gene c5167_047536, which belongs to the salutaridine reductase family.

Table 3. Hub gene promoter motif analysis (details about their E-value, width, Sites, and logos)

Nambe	Logo	E-value	Sites	Width
1		1.3e-005	5	50
2		1.0e-004	4	50
3		2.6e-003	4	50
4		7.3e-003	6	49
5		1.5e-001	4	40
6		1.3e+001	6	20

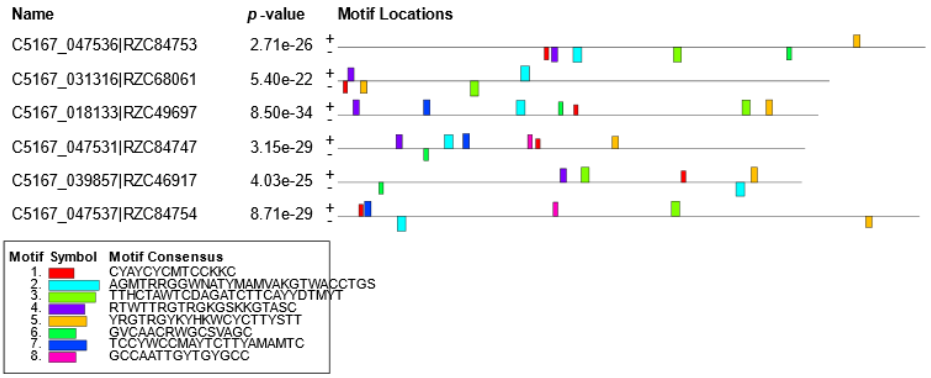


Figure 4. Hub gene location motifs (This diagram represents the locations of motif sites, with each block indicating the site's position and strength) (the highest p-value is linked to gene c5167-018133, which is associated with the Fe2OG dioxygenase domain-containing protein from the iron/ascorbate-dependent oxidoreductase family; in contrast, the lowest p-value is related to gene c5167_047536, which belongs to the salutaridine reductase family).

3.4. Investigation of the most important key enzymes in the morphine biosynthesis pathway

After entering the hub gene names into the KEGG mapper color tool, we generated Fig. 5. There are two pathways for morphine synthesis: the primary pathway and the secondary pathway. The primary pathway leads directly to the production of the alkaloid morphine, while the secondary pathway first produces codeine before morphine is formed. Eight genes are involved in the primary morphine production pathway, while nine genes are linked to the secondary pathway. The morphine biosynthesis process in the opium poppy includes several key genes and enzymes that are crucial for the production pathway (Hao, 2023). Key enzymes that play a role in the biosynthesis of morphine are codeine O-demethylase (PsCODM) and thebaine O-demethylase (PsT6ODM), which facilitate regioselective O-demethylation in the production of morphinans (Runguphan *et al.*, 2012). These enzymes are part of a group of six that are involved in converting (R)-reticuline into morphine, and they have been found in the sieve elements of the phloem (Runguphan *et al.*, 2012). Additionally, the final three enzymes that convert thebaine into morphine are among the most common active proteins present in latex. It is important to note that there are discrepancies concerning where morphine biosynthesis occurs. Although salutaridine synthesis appears to happen exclusively in sieve elements, most of the transformation from thebaine to morphine takes place in adjacent laticifers, which are abundant in latex-containing morphine (Runguphan *et al.*, 2012). The way biosynthetic processes are distributed in space may influence morphine production. In conclusion, improving morphine production likely requires optimizing the expression and activity of essential enzymes, including PsCODM and PsT6ODM, as well as those involved in the final steps of morphine synthesis. Genetic engineering methods, like CRISPR, could be used to alter these genes and possibly increase morphine output (Hao, 2023).

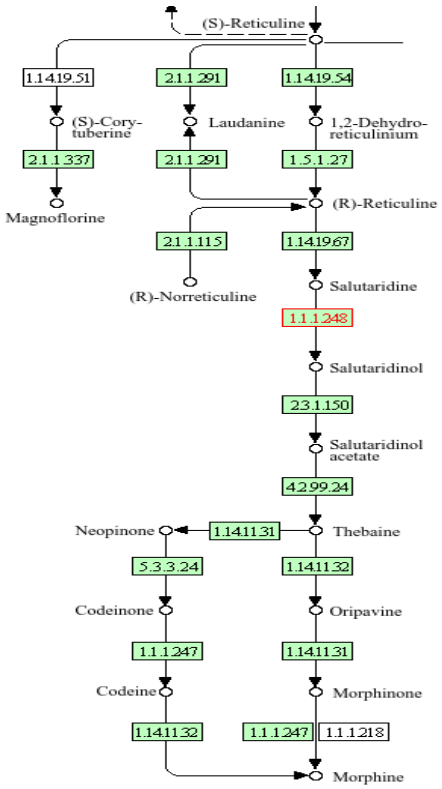


Figure 5. The alkaloid production pathways of the poppy plant include two main and secondary pathways (In the main pathway, morphine is directly metabolized to thebaine, orpavine, morphinone, and then morphine, and in the secondary pathway, thebaine is first converted to codeine and then to morphine).

3.5. Proposed pathway and the role of key genes in morphine biosynthesis

Considering the signaling role of hub genes, it is likely that these genes are located upstream of those

responsible for morphine production. The proteins they encode become activated in response to various internal and external stimuli that the plant encounters, including both biotic and abiotic stresses, triggering a series of downstream signaling cascades. Analyzing the sub-network genes associated with hub genes reveals that once the signaling pathway is activated by these hub genes, the resulting pathways activate genes that are vital for the synthesis of primary metabolites, especially those from the shikimate pathway. These pathways are crucial for generating precursors necessary for the biosynthesis of morphine and other metabolites like tyrosine. The analysis of promoter motifs in the hub genes revealed conserved motifs linked to their transcriptional regulation and interaction with transcription factors, as well as their involvement in membrane signal reception. These findings, along with the activities noted in the hub genes, reinforce the idea that these genes play a crucial role in sensing signals and triggering signaling responses mediated by transcription factors, which in turn affects the activation of primary and secondary response pathways in plants. In summary, plants respond to both biotic and abiotic signals by producing secondary metabolites, such as morphine.

4. Conclusion

This research sheds light on the intricate molecular environment that governs morphine production in *P. somniferum*. By analyzing protein–protein interactions (PPIs) and identifying key hub genes, a diverse range of participants involved in metabolic processes and redox regulation has been uncovered. The presence of both known and unknown proteins within the subnetwork underscores the complexity of the morphine biosynthetic pathway. The functional analysis of important hub genes, including those from the thioredoxin family, proteins with the Fe2OG dioxygenase domain, iron/ascorbate-dependent oxidoreductase family SCP domain proteins, and salutaridinol 7-O-acetyltransferase, highlights their varied functions. These genes are crucial in numerous processes, from redox regulation to various biological activities, underscoring their significance in shaping the dynamics of morphine biosynthesis. The gene ontology and pathway enrichment analyses provide a comprehensive understanding of the functional roles of the subnetwork. Important terms related to biological processes, cellular components, and molecular functions emphasize the involvement of subnetwork genes in the metabolic processes crucial for morphine production. Additionally, the KEGG pathway enrichment analysis strengthens this connection, highlighting the essential roles these genes have in metabolic pathways and the creation of secondary metabolites. This study significantly deepens our understanding of the molecular intricacies involved in morphine biosynthesis. The identified hub genes, enriched pathways, and regulatory elements establish a broad foundation for future research and genetic engineering initiatives aimed at enhancing the production of morphine and other alkaloids in opium poppy. The various functions of these genes open up possibilities for targeted approaches, offering chances to boost alkaloid yields and explore the broader biological significance of the morphine biosynthesis pathway.

Conflict of interests

All authors declare no conflict of interest.

Ethics approval and consent to participate

No humans or animals were used in the present research. The authors have adhered to ethical standards, including avoiding plagiarism, data fabrication, and double publication.

Consent for publications

All authors read and approved the final manuscript for publication.

Availability of data and material

All the data are embedded in the manuscript.

Authors' contributions

All authors had an equal role in study design, work, statistical analysis and manuscript writing.

Informed consent

The authors declare not to use any patients in this research.

Funding/Support

This study was supported by the Azarbaijan Shahid Madani University, Tabriz, Iran.

Acknowledgement

This article was achieved based on the material and equipment of Faculty of Agriculture, Azarbaijan Shahid Madani University, that the authors thanks it.

References

Aghaali Z., Naghavi M.R. 2024. Developing benzylisoquinoline alkaloid-enriched opium poppy via CRISPR-directed genome editing: A review. BMC Plant Biology 24(1): 700. https://doi.org/10.1186/s12870-024-05412-x

Aglas L., Soh W.T., Kraiem A., Wenger M., Brandstetter H., Ferreira F. 2020. Ligand binding of PR-10 proteins with a particular focus on the Bet v 1 allergen family. Current Allergy and Asthma Reports 20: 25. https://doi.org/10.1007/s11882-020-00918-4

Aleksander S.A., Balhoff J., Carbon S., Cherry J.M., Drabkin H.J., Ebert D., Feuermann M., Gaudet P., Harris N.L., Hill D.P., et al. 2023. The Gene Ontology knowledgebase in 2023. Genetics 224(1): iyad03. https://doi.org/10.1093/genetics/iyad031

Allen R.S., Miller J.A., Chitty J.A., Fist A.J., Gerlach W.L., Larkin P.J. 2008. Metabolic engineering of morphinan alkaloids by over-expression and RNAi suppression of salutaridinol 7-O-acetyltransferase in opium poppy. Plant Biotechnology Journal 6(1): 22-30. https://doi.org/10.1111/j.1467-7652.2007.00293.x

Aykanat S., Türktaş M. 2024. Divergent proteomic profiles of opium poppy cultivars. Turkish Journal of Biology 48(1): 80-90. https://doi.org/10.55730/1300-0152.2684

Bharadwaj R., Kumar S.R., Sharma A., Sathishkumar R. 2021. Plant metabolic gene clusters: Evolution, organization, and their applications in synthetic biology. Frontiers in Plant Science 12: 697318. https://doi.org/10.3389/fpls.2021.697318

Desgagné-Penix I., Farrow S.C., Cram D., Nowak J., Facchini P.J. 2012. Integration of deep transcript and targeted metabolite profiles for eight cultivars of opium poppy. Plant Molecular Biology 79: 295-313. https://doi.org/10.1007/s11103-012-9913-2

Dorafshan M., Soltani Howyzeh M., Shariati V. 2019. Identification of terpenoid backbone biosynthetic pathway genes in fruit of *Citrullus colocynthis* (L.) Schrad. medical plant by RNA sequencing. Iranian Journal of Medicinal and Aromatic Plants Research 35(4): 691-702. (In Farsi). https://doi.org/10.22092/ijmapr.2019.125294.2510

Fazili M.A., Bashir I., Ahmad M., Yaqoob U., Geelani S.N. 2022. In vitro strategies for the enhancement of secondary metabolite production in plants: A review. Bulletin of the National Research Centre 46(1): 35. https://doi.org/10.1186/s42269-022-00717-z

Filiault D.L., Ballerini E.S., Mandáková T., Aköz G., Derieg N.J., Schmutz J., Jenkins J., Grimwood J., Shu S., Hayes R.D. 2018. The *Aquilegia* genome provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a unique history. eLife 7: e36426. https://doi.org/10.7554/eLife.36426

Garg A., Agrawal L., Misra R.C., Sharma S., Ghosh S. 2015. *Andrographis paniculata* transcriptome provides molecular insights into tissue-specific accumulation of medicinal diterpenes. BMC Genomics 16: 659. https://doi.org/10.1186/s12864-015-1864-y

Ghorbani A., Rostami M., Izadpanah K. 2023. Gene network modeling and pathway analysis of maize transcriptomes in response to Maize Iranian mosaic virus. Genomics 115(3): 110618. https://doi.org/10.1016/j.ygeno.2023.110618

Grutsch S., Fuchs J.E., Ahammer L., Kamenik A.S., Liedl K.R., Tollinger M. 2017. Conformational flexibility differentiates naturally occurring Bet v 1 isoforms. International Journal of Molecular Sciences 18(6): 1192. https://doi.org/10.3390/ijms18061192

Hao L. 2023. Evaluation of biosynthetic pathway and engineered biosynthesis of morphine with CRISPR. 5th International Conference on Biotechnology and Biomedicine 59: 01022. https://doi.org/10.1051/bioconf/20235901022

Higashi Y., Kutchan T.M., Smith T.J. 2011. Atomic structure of salutaridine reductase from the opium poppy (*Papaver somniferum*). Journal of Biological Chemistry 286(8): 6532-6541. https://doi.org/10.1074/jbc.M110.168633

Huacachino A.A., Joo J., Narayanan N., Tehim A., Himes B.E., Penning T.M. 2024. Aldo-keto reductase (AKR) superfamily website and database: An update. Chemico-Biological Interactions 398: 111111. https://doi.org/10.1016/j.cbi.2024.111111

Jha K., Saha S., Dutta P. 2024. Incorporation of gene ontology in identification of protein interactions from biomedical corpus: A multi-modal approach. Annals of Operations Research 339(3): 1793-1811.

Karimizadeh E., Sharifi-Zarchi A., Nikaein H., Salehi S., Salamatian B., Elmi N., Gharibdoost F., Mahmoudi M. 2019. Analysis of gene expression profiles and protein-protein interaction networks in multiple tissues of systemic sclerosis. BMC Medical Genomics 12: 199. https://doi.org/10.1186/s12920-019-0632-2

Krishnamurthy P., Pothiraj R., Suthanthiram B., Somasundaram S.M., Subbaraya U. 2022. Phylogenomic classification and synteny network analyses deciphered the evolutionary landscape of aldo–keto reductase (AKR) gene superfamily in the plant kingdom. Gene 816: 146169. https://doi.org/10.1016/j.gene.2021.146169

Kuiper C., Vissers M.C. 2014. Ascorbate as a co-factor for Fe-and 2-oxoglutarate dependent dioxygenases: Physiological activity in tumor growth and progression. Frontiers in Oncology 4: 359. https://doi.org/10.3389/fonc.2014.00359

Li M., Li D., Tang Y., Wu F., Wang J. 2017. CytoCluster: A cytoscape plugin for cluster analysis and visualization of biological networks. International Journal of Molecular Sciences 18(9): 1880. https://doi.org/10.3390/ijms18091880

Liu X., Hong Z., Liu J., Lin Y., Rodríguez-Patón A., Zou Q., Zeng X. 2020. Computational methods for identifying the critical nodes in biological networks. Briefings in Bioinformatics 21(2): 486-497. https://doi.org/10.1093/bib/bbz011

Malik C., Dwivedi S., Rabuma T., Kumar R., Singh N., Kumar A., Yogi R., Chhokar V. 2023. De novo sequencing, assembly, and characterization of *Asparagus* racemosus transcriptome and analysis of expression profile of genes involved in the flavonoid biosynthesis pathway. Frontiers in Genetics 14: 1236517. https://doi.org/10.3389/fgene.2023.1236517

Manni M., Berkeley M.R., Seppey M., Simão F.A., Zdobnov E.M. 2021. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Molecular Biology and Evolution 38(10): 4647-4654. https://doi.org/10.1093/molbev/msab199

Mogensen J.E., Wimmer R., Larsen J.N., Spangfort M.D., Otzen D.E. 2002. The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. Journal of Biological Chemistry 277(26): 23684-23692. https://doi.org/10.1074/jbc.M202065200

Osorio-Concepción M., Lax C., Navarro E., Nicolás F.E., Garre V. 2021. DNA methylation on N6-adenine regulates the hyphal development during dimorphism in the early-diverging fungus *Mucor lusitanicus*. Journal of Fungi 7(9): 738. https://doi.org/10.3390/jof7090738

Ozber N., Yu L., Hagel J.M., Facchini P.J. 2023. Strong feedback inhibition of key enzymes in the morphine biosynthetic pathway from opium poppy detectable in engineered yeast. ACS Chemical Biology 18(2): 419-430. https://doi.org/10.1021/acschembio.2c00873

Radauer C., Lackner P., Breiteneder H. 2008. The Bet v 1 fold: An ancient, versatile scaffold for binding of large, hydrophobic ligands. BMC Evolutionary Biology 8: 286. https://doi.org/10.1186/1471-2148-8-286

Runguphan W., Glenn W.S., O'Connor S.E. 2012. Redesign of a dioxygenase in morphine biosynthesis. Chemistry & Biology 19(6): 674-678. https://doi.org/10.1016/j.chembiol.2012.04.017

Shanbhag A.P., Bhowmik P. 2024. Cancer to cataracts: The mechanistic impact of aldo-keto reductases in chronic diseases. The Yale Journal of Biology and Medicine 97(2): 179. https://doi.org/10.59249/VTBV6559

Tahmasebi A., Ebrahimie E., Pakniyat H., Ebrahimi M., Mohammadi-Dehcheshmeh M. 2019. Tissue-specific transcriptional biomarkers in medicinal plants: Application of large-scale meta-analysis and computational systems biology. Gene 691: 114-124. https://doi.org/10.1016/j.gene.2018.12.056

Tolchin Z.A., Dukes D.M., Gharbaoui L.M., Smith J.M. 2023. Dearomative access to (−)-thebaine and derivatives. Organic Letters 25(47): 8424-8428. https://doi.org/10.1021/acs.orglett.3c03270

van Dorp E.L., Yassen A., Dahan A. 2007. Naloxone treatment in opioid addiction: The risks and benefits. Expert Opinion on Drug Safety 6(2): 125-132. https://doi.org/10.1517/14740338.6.2.125

Vorontsov I.E., Kozin I., Abramov S., Boytsov A., Jolma A., Albu M., Ambrosini G., Faltejskova K., Gralak A.J., Gryzunov N., et al. 2024. Cross-platform DNA motif discovery and benchmarking to explore binding specificities of poorly studied human transcription factors. bioRxiv. https://doi.org/10.1101/2024.11.11.619379

Yucebilgili Kurtoglu K., Unver T. 2021. Integrated omics analysis of benzylisoquinoline alkaloid (BIA) metabolism in opium poppy (*papaver somniferum* L.). In: Tombuloglu H., Unver T., Tombuloglu G., Hakeem K.R. (eds) Oil Crop Genomics. Springer, Cham. https://doi.org/10.1007/978-3-030-70420-9_13

HOW TO CITE THIS ARTICLE

Bagheri N., Tarinejad A., Hasanpur K., Majidi M. 202x. Decoding the Biosynthetic Pathway of the Alkaloid Morphine with Bioinformatics. *Agrotechniques in Industrial Crops* x(x): xx-xx. 10.22126/ATIC.2025.11523.1182