Evaluation of Biosynthetic Pathway and Engineered Biosynthesis of Morphine with CRISPR

Luyang Hao[a]

College of Arts and Sciences, The Ohio State University, Columbus, Ohio, 43212, United States

Abstract: As a secondary metabolite produced by opium poppy (*Papaver somniferum*), morphine is a valuable subject for research and application purposes within the medical field. Thanks to the research conducted throughout the past few decades, the main pathway for the biosynthesis of morphine has been thoroughly elucidated. In addition, whole genome sequencing and evolutionary genomics of opium poppy have provided us with crucial information regarding specific details of such a pathway on the molecular level, which includes coding regions and functions of key enzymes that play vital roles in the production of specific metabolites. These advantages can be combined with the application of current genome editing techniques, such as CRISPR, to allow the regulated and optimized production of desirable metabolites through manipulating particular genes. This research focuses on the functional evaluation of key enzymes in the biosynthetic pathway of morphine in opium poppy, further exploring possibilities of regulated production of morphine with CRISPR.

1 Introduction

Opium poppy (*Papaver somniferum*) is medicinally valuable due to its production of several benzylisoquinoline alkaloids (BIAs) of significant pharmaceutical importance. Derivatives such as morphine and codeine are categorized as narcotic analgesics and are widely used in the medical field for pain relief or cough suppression [1]. Despite its addictive nature and other side effects, morphine is an effective substance that can aid patients afflicted with both acute and chronic pain, which justified its adoption in many developed countries that have imposed strict regulations on its usage. Currently, morphine is more commonly applied to cancer patients or those who require end-of-life care. [13]. Consequently, as cancer quickly becomes a more prevalent threat accompanied by the trend of global population aging, the demand for morphine has vastly increased to supply the need for potent analgesics. Moreover, there is also a potential need for morphine to fulfill its need in scientific research. The addictive nature of morphine is an exciting aspect studied to understand the psychological foundations of addiction, which can be expanded in behavioral neuroscience and psychology studies.

The modern method of morphine production involves direct extraction from opium poppy, and thus the efficiency of such extraction would dictate the amount of morphine produced. One ideal approach to elevate the efficiency would be to fundamentally regulate the biosynthetic pathway of morphine within opium poppy so that the desired metabolites can be generated in optimal quantities in a controlled manner. This approach requires modifications on key enzymes throughout the pathway that essentially lead back to the genes that encode critical information regarding enzymatic functions. However, regulation of the metabolic pathway of morphine and other associating intermediates or products has not been completely elucidated yet, which becomes the remaining issue. In our study, we have analyzed and cumulated the findings and results from others to seek a viable solution to this problem. The evaluation would consist of three main aspects, which include the molecular and genetic levels of the biosynthetic pathway of morphine as well as the discussion on clustered regularly interspaced short palindromic repeats (CRISPR) and its application in the regulation of morphine biosynthesis.

2 Biosynthetic Pathway of Morphine

The standard biosynthesis of morphine is a multi-step mechanism consisting of a collection of vital enzymes that stimulate the reactions.

2.1 General Pathway

As shown in Figure 1, the biosynthesis of morphine begins with the condensation of dopamine and 4-hydroxyphenylacetaldehyde by norcoclaurine synthase (NCS), yielding (*S*)-norcoclaurine as an intermediate. A series of consecutive reactions involving norcoclaurine 6-*O*-methyltransferase (6OMT), (*S*)-coclaurine *N*-methyltransferase (CNMT), *N*-methylcoclaurine 3'-hydroxylase (NMCH), and 3'-hydroxy *N*-methylcoclaurine 4'-*O*-methyltransferase (4'OMT) would

[a] hly119477992@gmail.com

then take place to transform (*S*)-norcolaurine into (*S*)-reticuline, which is a major intermediate and a critical branch point. Many derivatives can be formed from (*S*)-reticuline. Specifically for synthesizing morphine, (*S*)-reticuline would be epimerized to form (*R*)-reticuline. Next, salutaridine would be formed via the catalyzation of salutaridine synthase through intramolecular carbon-carbon phenol coupling. Notably, salutaridine is the first tetracyclic promorphinan alkaloid formed. Subsequently, salutaridine reductase (SalR) would reduce the C7 keto group of salutaridine in a stereospecific manner to yield salutaridinol. This process is known to be NADPH-dependent. Salutaridinol is then catalyzed by 7-*O*-acetyltransferase to form salutaridinol 7-*O*-acetate, which would simultaneously lose the acetyl group and rearrange into thebaine. Thebaine is another key intermediate that would be *O*-demethylated by thebaine 6-*O*-demethylase (T6ODM), resulting in the formation of neopinone. Neopinone is then catalyzed by neopinone isomerase (NISO) to transform into codeinone [1, 3]. With NADPH present, a reduction event stimulated by codeinone reductase (COR) would take place to convert codeinone to codeine as one of the major products of the pathway. Finally, morphine would be formed with the demethylation of codeine by codeine-*O*-demethylase (CODM) [1].

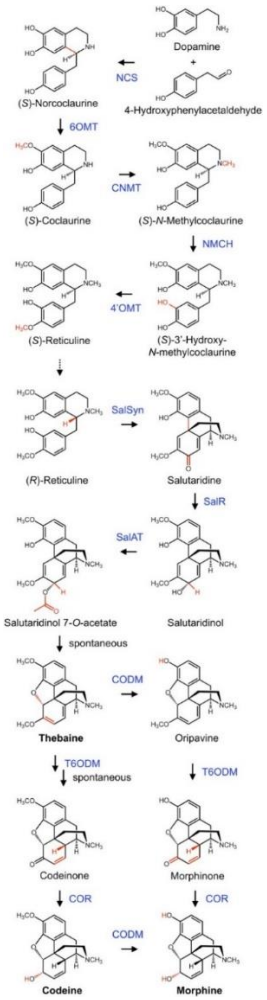

Figure 1. Biosynthesis of Morphine in Opium Poppy from the Tyr Derivatives Dopamine and 4-Hydroxyphenylacetaldehyde [1].

2.2 Notable Enzymes within Pathway

Within the biosynthetic pathway, a few crucial enzymes that determine the production of several different metabolites. For instance, SalR is a member of the short-chain dehydrogenase/reductase family that is vital to reducing the C-7 keto group of salutaridine, which ultimately decides the production of morphinan alkaloids. Its atomic structure and genetic position within the genome of opium poppy have been thoroughly assessed [5]. Another enzyme worth mentioning is NISO. It was three years ago that the importance of NISO was understood. A study conducted in 2019 emphasized the role NISO played in the isomerization of neopinone to codeinone which was previously assumed to be a spontaneous reaction[4]. This finding relied on the integration of heterologous expression of opiate biosynthetic enzymes in yeast cells [3]. T6ODM is another enzyme that is widely studied in the pathway. It is known to facilitate O-demethylation in regioselective manners. Thus, previous studies have focused on the specific binding sites of such enzyme and the overall mechanism of the reaction it catalyzes [6].

3 Genome of Opium Poppy and Applications of CRISPR in Genomic Research

3.1 Genome of Opium Poppy

Although the biosynthetic pathway of opium poppy has been sophisticatedly studied in the past few decades, the genome of the primary plant involved in such pathway had remained unknown until the reference genome was first reported in 2018 with the adoption of whole genome sequencing that was used to decode gene content that essentially leads to the biosynthetic pathway of morphine. The findings have revealed a complex genome with 70% of repetitive sequences. 51213 protein-coding genes were assessed along with 9494 non-coding RNAs. PacBio sequencing and Oxford Nanopore Technologies (ONT) were the two unique technologies the researchers adopted to achieve longer reads that guarantees the elongation of scaffolds spanning repeat-rich genomic regions, ensuring quality control. A 25-kb missing portion of the highly repetitive genomic region of 227 kb was corrected, corresponding to the T6ODM gene family that codes for T6ODM [2]. This approach was quite insightful in providing an effective method to decode other complex plant genes, including other portions of the opium poppy genome. Previously, a relationship between clusters of functionally related genes and the biosynthesis of specific metabolites has been established. Whole genome sequencing has therefore paved the way for scientists to explore and determine the locations of all related genes involved in BIA biosynthesis and investigate their evolutionary history. Two remarkable findings include the whole genome duplication of opium poppy around 7 to 8 million years ago and the STORR gene fusion event. As the product resulted from a fusion event in the ancestral genome duplication, the STORR gene was acknowledged

as the key to understanding the biosynthesis of morphinan in opium poppy [2]. Most importantly, the information provided by whole genome sequencing dramatically assists us in studying the functions of genes by reverse genetics, and one of the tools for us to manipulate gene expression is the CRISPR/Cas system.

3.2 CRISPR/Cas System

Genome editing has been a controversial yet innovative field of study in biological sciences in recent years that aims to recognize and locate specific sequences within an organism's genome accurately and precisely and ultimately edit its genome for various purposes. To this day, the development of gene-editing techniques has allowed scientists to adopt several advanced techniques to accomplish genome editing. Clustered regularly interspaced short palindromic repeats, or CRISPR, was one of them. First discovered in 1987 in the defense mechanism of *E. coli* cells, it has come to be, if not, one of the most powerful and productive tools for genome editing nowadays [7].

3.2.1 Basic Mechanism of CRISPR.

CRISPR is an extensive class of short palindromic repeat sequences within a wide range of prokaryotes complementary with some foreign DNAs, including those of viruses. Infection of viruses would trigger the generation of such DNA to bind to the viral ones. Nuclease categorized as Cas would then cut the viral DNA into pieces for self-defense. This acquired immune mechanism is a genome-editing tool with great potential given by mother nature [7]. It was not until over two decades after the original discovery of the CRISPR system that scientists could reconstruct such a system *in vitro* that possesses basic functionalities or is capable of cutting individual DNA sequences.

The mechanism of the CRISPR/Cas system revolves around two nuclease domains named RuvC and HNH, accompanied by a PAM-interacting domain (PI) (Fig. 2). The protospacer adjacent motif (PAM) would direct the binding of Cas at specific locations of the DNA strand to allow the cleaving of the double-stranded DNA by RuvC and HNH domains and the formation of a double-strand break (DSB). The break would then be repaired by the natural repair mechanisms of the cell, which can be generalized into two pathways, non-homologous end joining repair (NHEJ) and homology-directed repair (HDR) (Fig. 3). As shown in Figure 3, the HDR usually requires a template DNA that is highly related to the DSB. The requirements of the Cas nuclease, as well as the PAM sequence, have ensured the accuracy and specificity of the CRISPR/Cas system. Meanwhile, the development of various CRISPR types due to the differences in the Cas enzyme has given rise to many utilities in genome editing and applied biological research [7].

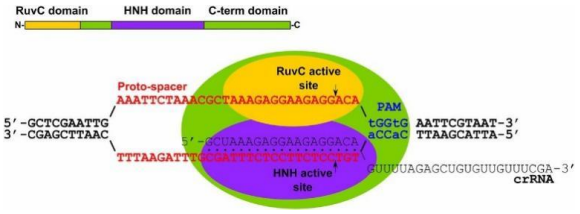

Figure 2. General structure of the Cas enzyme. HNH cleaves DNAs complementary to the RNA guide, while RuvC cleaves the non-complementary ones. The prerequisites for both pathways (NHEJ and HDR) are the protospacer adjacent motif (PAM) [7, 11]

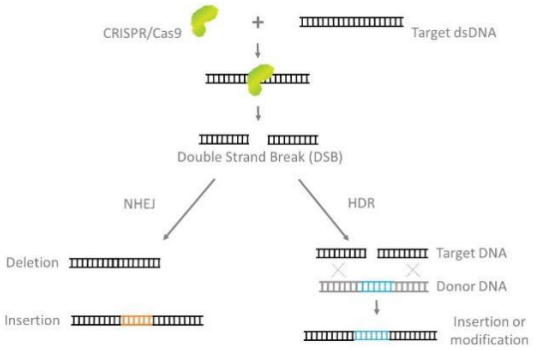

Figure 3. Basic mechanism of CRISPR/Cas system and utilities. CRISPR accomplishes genome editing in two major pathways, gene knockout and gene knock-in. HDR requires a template donor DNA strand to accomplish gene knockin, as shown above [7, 12].

3.2.2 Silencing.

Currently, most of the research conducted with CRISPR/Cas has focused on silencing individual genes through gene knockout, which heavily relies on the fundamental repair mechanisms of cells. When a double-stranded DNA is cut, a DSB is generated and repaired by NHEJ in most cases [7]. Naturally, the deletion or addition of one or more nucleotides would inevitably interrupt the expression of the target gene because the base frameshift has occurred at which a protein-coding gene was previously present, hence the outcome of gene silencing or knockout.

3.2.3 Overexpression.

Overexpression is another aspect of the application regarding the CRISPR/Cas system. Compared to agrobacterium-mediated gene transformation traditionally used to overexpress individual genes through T-DNA insertion, CRISPR is more effective in avoiding many side effects caused by random insertions, which include the silencing of other functional genes [7]. After cutting the DNA sequence at a specific site, HDR would insert a foreign sequence into the cleavage site while ensuring the impacts caused by the interruption of other genes are minimal, considering that a variety of genes are indeed position-specific.

3.2.4 Advantages and Limitations.

Today, the CRISPR/Cas system is highly regarded as one of the most efficient and user-friendly gene editing tools in biological research due to its compatibility and robustness. Compared to conventional tools, CRISPR has emphasized the accuracy and specificity of its editing site, which effectively minimized the impact brought by position-specific genes and reduced the issues associated with uncertainty due to random deletions or insertions that have discombobulated scientists in the past. Moreover, the variance among the Cas enzymes has quickly been utilized by scientists to develop a collection of useful tools for specific purposes. The utility of CRISPR is not limited to silencing or overexpressing a specific gene; gene replacement and CRISPR/Cas mediated base editing/regulation and epigenome editing also have excellent research values and are prospectively beneficial to the perfection in genomic edition [7]. Nevertheless, CRISPR also suffers from setbacks due to its physical limitations. Since a frameshift would unavoidably happen in the DNA, the knockout of a protein-coding gene would often cause nonsense mutations or sometimes code for a different protein. On the other hand, silencing small RNAs can be quite challenging to accomplish and is often ineffective because base frameshift has little effect on targeting a miRNA on a protein-coding gene. Besides, there are difficulties associated with the gene knock-in using CRISPR. The insertion or replacement of long sequences and genes is difficult to achieve and often results in low efficiency [7].

3.2.5 Engineered Biosynthesis of Morphine with CRISPR.

With the elucidation of the biosynthetic pathway of morphine in opium poppy and the discovery of key metabolic enzymes, genome editing is now accessible for scientists to explore possibilities to modify the production of morphine through genome editing. Generally speaking, the functional verification of the pathway is achieved through the partial or complete inhibition of enzymatic functions by gene deletion or expression silencing. Changes in the metabolites were assessed to determine the functions of candidate pathway genes [8]. Previous efforts to silence single genes in the biosynthetic pathway have solidified this idea. For instance, T6ODM, as mentioned previously, converts thebaine to neopinone. The knockdown of it would result decrease the accumulation of codeine and morphine and increase that of thebaine as it becomes the final product due to a lack of functional enzymes. Similarly, the silence of (*S*)-Reticuline would terminate the pathway at an even early stage, reducing the production of thebaine, papaverine, and codeine altogether. On the contrary, the overexpression of the same gene would directly increase the production in the final products, including codeine and morphine, suggesting that (*S*)-reticuline is a crucial component to the pathway and could potentially be modified on the genetic level to increase the production of morphine or thebaine [10]. Based on these previous outcomes, A study conducted in 2018 demonstrated the mechanism in an attempt to simultaneously overexpress and silence multiple genes associated with the biosynthesis of BIAs in opium poppy. With the adoption of a simultaneous silencing construct, three individual genes were knocked down simultaneously, and significant changes in the accumulation of metabolites were found. It was observed that the suppression of T6ODM, CODM, and BBE1 genes that lead to the silencing of the corresponding enzymes have led to a notable increase in thebaine and papaverine and a decrease in codeine, morphine**,** sanguinarine, and noscapine [10]. Although the exact mechanism was applied and the difference between the silencing of singular and multiple genes was little, this approach was a great approach to rule out some potential factors hidden within some unknown pathways. On the other hand, it also gave us a new perspective on our goal of regulating morphine synthesis.

In another study, the effectiveness of the CRISPR/Cas system in the genome editing of opium poppy was successfully demonstrated. Researchers used synthetic and viral plasmid backbones to knock out the 4'OMT gene that regulates the synthesis of multiple aromatic metabolites, including noscapine and morphine. This study was conducted on the foundation of a previous approach that adopted a tobacco rattle virus (TRV) based post-transcriptional gene on achieving gene silencing in opium poppy. The CRISPR/Cas9 system was modified and improved by using the TRV vector to express single guide RNAs (sgRNAs). The results were significant in that the target metabolite accumulation level was significantly lower because of the interrupted biosynthetic pathway due to gene knockout. One notable insight is that the CRISPR/Cas system has shown high consistency and efficiency in gene knockout. These advantages have favored using the CRISPR/Cas system for gene editing in opium poppy, and this study can serve as a framework for future references regarding morphine biosynthesis [9].

4 Conclusion

As the plant known to produce narcotic analgesics morphine and codeine, opium poppy has always remained a subject of interest in modern medicine [1]. With the biosynthetic pathway of morphine elucidated, we can locate and observe critical enzymes that dictate the production of BIAs, followed by the adoption of whole genome sequencing that took a step further to relate the enzymes to the genome, including their respective coding regions as well as evolutionary history. These achievements have allowed us to evaluate the functionalities of enzymes on a genetic level by the CRISPR/Cas system. This advanced gene-editing tool allows us to explore innovative ways to manipulate the production of morphine and other BIAs by gene silencing or overexpression. On balance, we have concluded that the biosynthetic pathway of morphine would be subject to further modifications for maximized production, which could make significant contributions to the medicinal field when it comes to real-world applications.

References

1 Onoyovwe, A., Hagel, J., Chen, X., Khan., M. F., Schriemer, D. C., Facchini, P. J., 2013. Morphine Biosynthesis in Opium Poppy Involves Two Cell Types: Sieve Elements and Laticifers. The Plant Cell, 25(10): 4110–4122.

2 Hu, Y., Zhao, R., Xu, P., Jiao, Y., 2018. The genome of Opium Poppy reveals evolutionary history of Morphinan pathway. Genomics, Proteomics & Bioinformatics, 16(6): 460–462.

3 Dastmalchi, M., Chen, X., Hagel, J. M., Chang, L., Chen, R., Ramasamy, S., Yeaman, S., Facchini, P. J., 2019. Neopinone isomerase is involved in codeine and morphine biosynthesis in opium poppy. Nature Chemical Biology, 15(4): 384–390.

4 Sohrabi, S. M., Ismaili, A., Nazarian-Firouzabadi, F., 2018. Simultaneous over-expression and silencing of some benzylisoquinoline alkaloid biosynthetic genes in opium poppy. Industrial Crops and Products, 123: 581–590.

5 Ziegler, J., Facchini, P. J., Geißler, R., Schmidt, J., Ammer, C., Kramell, R., Voigtländer, S., Gesell, A., Pienkny, S., Brandt, W., 2009. Evolution of morphine biosynthesis in opium poppy. Phytochemistry, 70(15-16): 1696–1707.

6 Wijekoon, C. P., Facchini, P. J., 2011. Systematic knockdown of morphine pathway enzymes in opium poppy using virus-induced gene silencing. The Plant Journal, 69(6): 1052–1063.

7 Zhang, D., Zhang Z., Unver, T., Zhang, B., 2021. CRISPR/Cas: A powerful tool for gene function study and crop improvement. Journal of Advanced Research, 29: 207-221.

8 Zhu, X., Liu, X., Liu, T., Wang, Y., Ahmed, N., Li, Z., Jiang, H., 2021. Synthetic Biology of Plant Natural Products: From pathway elucidation to engineered biosynthesis in plant cells. Plant Communications, 2(5): 100229.

9 Alagoz, Y., Gurbok, T., Zhang, B., Unver, T., 2016. Manipulating the Biosynthesis of Bioactive Compound Alkaloids for Next-Generation Metabolic Engineering in Opium Poppy Using CRISPR-Cas 9 Genome Editing Technology. Scientific Reports, 6: 30910.

10 Sohrabi, S. M., Ismaili, A., Nazarian-Firouzabadi, F., 2018. Simultaneous over-expression and silencing of some benzylisoquinoline alkaloid biosynthetic genes in opium poppy. Industrial Crops and Products, 123:581-590.

11 Gasiunas, G., Barrangou, R., Horvath P., Siksnys, V., 2012. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences of the United States of America, 109 (39): E2579-E2586

12 Ruiz, F. CRISPR: Discovery and Adaptation. 2020. Gold Biotechnology. https://www.goldbio.com/articles/article/CRISPR-discovery-and-adaptation

13 U.S. National Library of Medicine. 2022. Morphine: Medlineplus drug information. https://medlineplus.gov/druginfo/meds/a682133.html.