METABOLIC PATHWAY RECONSTRUCTION

OF Hevea brasiliensis

KHOO BEE FEONG

UNIVERSITI SAINS MALAYSIA

2011

METABOLIC PATHWAY RECONSTRUCTION OF *Hevea brasiliensis*

by

KHOO BEE FEONG

Thesis submitted in fulfillment of the requirements
for the degree of
Master of Science

April 2011

Dedicated to the truthness in you

ACKNOWLEDGEMENT

I would like to thank my family members for being around and apologize for all the problems I had caused in the period of the project. I would like to thank my supervisor, Prof Maqsudul Alam for all the opportunities and constructive advices that I ever got. Lastly, I would like to thank all the CCB@USM members who walk through the journey with me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AraCyc	A publicly accessible database of *Arabidopsis thaliana*
ER	Endoplasmic reticulum
GO	Gene ontology
ha	Hectares
IgE	Immunoglobulin E
H. brasiliensis	*Hevea brasiliensis*
KEGG	Kyoto Encyclopedia of Genes and Genomes
NCBI	National Center for Biotechnology Information
PoplarCyc	A publicly accessible database of *Populus trichocarpa*

Pembinaan semula laluan metabolik *Hevea brasiliensis*

ABSTRAK

Hevea brasiliensis merupakan tumbuhan berkayu yang menghasilkan getah. Kajian selama satu dekad telah dijalankan atas pokok ini. Namun laluan metabolismanya yang memuaskan masih belum dibina. Projek ini adalah mengenai pembinaan laluan metabolic dan kurasi manual kepada laluan yang berkaitan dengan penghasilan susu getah. Laluan-laluan ini dibina dan disah berdasarkan data genomic dan transkriptomik. Thesis ini akan membincangkan tentang pembinaan semula laluan metabolik, laluan dan jaringan yang berkaitan dengan penghasilan susu getah oleh *H. brasiliensis* yang dikurasi secara manual. Lima puluh sela telah diisi secara manual ke dalam 65 laluan metabolic. Thesis ini mengandungi 15 laluan metabolik yang dibina semula, 4 laluan dan 2 jaringan yang yang dikurasikan secara manual.

Metabolic pathway reconstruction of *Hevea brasiliensis*

ABSTRACT

Hevea brasiliensis is a latex-producing woody plant. This plant has been studied for more than one decade. However, its metabolic pathways are still not well established. This project is about reconstruction of the metabolic pathways and manual construction of pathways related to latex production. The pathways were reconstructed and validated based on genomic and transcriptomic data. This thesis will discuss about the metabolic pathways reconstructed as well as the manually curated pathways and networks of *H. brasiliensis* related to latex production. Fifty gaps were manually closed in 65 reconstructed metabolic pathways. This thesis included 15 reconstructed metabolic pathways, 4 manually curated pathways, and 2 manually curated networks.

CHAPTER 1

INTRODUCTION

Natural rubber is a biodegradable polymer synthesized by at least 2,000 species of plants belonging to 300 genera. *H. brasiliensis* is one of the important industrial crops for natural rubber production. Throughout the years, most of the research on *H. brasiliensis* had been focused on the latex production. Latex is the cytoplasm of the specialized sieve tube called laticifer. The latex is produced in the laticifer as the plant grows up. With the advances in genome sequencing technology, the *H. brasiliensis* genome has been decoded.

The huge amount of genome and transcriptome data generated prompt for a better visualization, validation and utilization of the data. This project is about metabolic pathway reconstruction, manual curation of pathways, and network analysis to understand the influence of different cellular activities on latex production. *In silico* studies are vital in our understanding of metabolic activity, disease resistance and to further improve the latex production of *H. brasiliensis*.

CHAPTER 2

LITERATURE REVIEW

2.1 Rubber tree

Rubber tree is an important industrial crop for natural rubber production. At present, more than 9.5 million hectares (ha) in about 40 countries are devoted to rubber tree cultivation with a production of about 6.5 million tons of dry rubber each year (Venkatachalam *et al*., 2006). The world supply of natural rubber is barely keeping up with a global demand for 12 million tons of natural rubber in 2020 (Venkatachalam *et al*., 2006). Due to the increasing global demand on rubber, further improvement of natural rubber production is necessary (Ziegler *et al*., 2009). Rubber plantations are expanding rapidly throughout montane mainland South East Asia. More than 500,000 ha may have been converted already in the uplands of China, Laos, Thailand, Vietnam, Cambodia, and Myanmar (Ziegler *et al*., 2009). It is estimated that by year 2050, the area of land dedicated to rubber and other diversified farming systems could more than double or triple (Ziegler *et al*., 2009).

2.2 Natural Rubber

Natural rubber is synthesized by at least 2,000 species of plants belonging to 300 genera. The rubber polymerases from different plants produce natural rubber with different quality. Some of the research done on non-*Hevea* plant species able to produce natural rubber include Dandelion (van Beilen and Poirier, 2007), *Taraxacum* (Wahler *et al*., 2009; Buranov *et al*., 2010), and *Euphorbia lactiflua* (Cornish *et al*., 1999; Luigi Barbieri *et al*., 1983; Wood *et al*., 2000). However, *H. brasiliensis* is the only economically viable source of natural rubber due to its good yield of rubber and

the excellent physical properties of the rubber products. *Parthenium argentatum* is an alternative to *H. brasiliensis* due to its low cost production of natural rubber (Kim *et al*., 2004; Rao *et al*., 2011).

2.3 Properties and applications of natural rubber

Latex is a biodegradable polymer. Throughout the years, numerous isolates able to degrade latex were identified (Yikmis *et al*., 2008). This special ability has been proposed as the chemotaxonomy (Linos *et al*., 2002). The isolates able to degrade natural rubber include *Actinoplanes*, *Streptomyces*, *Gordonia*, *Mycobacterium*, and *Nocardia* (Warneke *et al*., 2007)

Natural rubber is widely used worldwide because of its excellent properties in terms of high elasticity and mechanical strength, including high resistance to impact and tear as well as low heat build-up during deformation (Intharapat *et al*., 2009). An example of a natural rubber application is the production of composite. Composite of natural rubber and other synthetic compounds are produced to increase the mechanical strength of the material (Ochigbo *et al*., 2009).

The application of concrete-polymer composites, especially polymer-modified cement mortar and concrete, is expanding rapidly because of their good properties compared to conventional cement mortar and concrete. A study on the characterization of SAE latex in mortar showed that the SAE latex has air entrainment effect, increasing the air content and reducing the bulk density of the fresh mortar. The SAE latex influences the development of the compressive strength but slightly on the flexural strength, and improves the toughness, shrinkage property, waterproofing quality, and anti-penetration capacity of the mortar significantly (Wang and Wang, 2010).

2.4 Morphology of laticifer

Throughout the years, most of the research on *H. brasiliensis* has been focused on the latex yield (Sando *et al*., 2009). Latex is the cytoplasm of the specialized sieve tube called laticifer (Sando *et al*., 2009). When the bark is tapped, the cytoplasmic contents of these laticifers are expelled in the form of latex (Sando *et al*., 2009). Morphological study of *H. brasiliensis* reveal the structure of articulated and non-articulated laticifer (Sando *et al*., 2009). The histological study of the laticifer using confocal microscopy shows the distribution of morphologically different laticifer cells at different layers of the bark of *H. brasiliensis* (Sando *et al*., 2009). This study confirmed the presence of concentric-layered articulated laticifers in secondary phloem in the bark (Figure 2.1) (Sando *et al*., 2009). It also shows the occurrence of non-articulated laticifers at the outermost laticifer layer in the bark (Figure 2.2) (Sando *et al*., 2009).

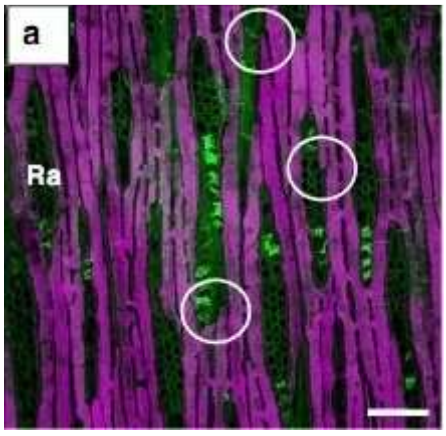

Figure 2.1 Micrograph of articulated laticifer (purple), circled articulated part of the laticifer, scale bar 100 µm (Figure from Sando *et al*., 2009; Figure 3(a); page 221)

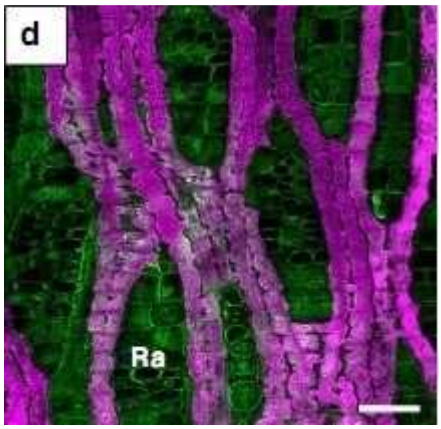

Figure 2.2 Micrograph of non-articulated laticifer (purple), scale bar 100 µm (Figure from Sando *et al*., 2009; Figure 3(d); page 221)

Laticifers at an early stage of development were welly distinguished by the presence of the nucleus and dense cytoplasm rich in ribosomes, small vacuoles, mitochondria, endoplasmic reticulum, dictyosomes, and osmiophilic bodies (Cai *et al*., 2009). The double layer plasma membrane was also visible under the microscope (Cai *et al*., 2009). During the vacuolization stage, small vacuoles containing rubber particles were formed from the ER (Cai *et al*., 2009). The resultant small vacuoles enlarged and began to accumulate latex particles (Cai *et al*., 2009). At last, these vacuoles became electron transparent and contained latex particles of various sizes which were later released into a central vacuole (Cai *et al*., 2009). At maturity, laticifer cells consist of a thin layer of cytoplasm, central large vacuole filled with rubber particles, very thick cell wall, and plasmodesmata are not observed between laticifers and the neighbor cells (Cai *et al*., 2009).

2.5 Latex biosynthesis under ethylene stimulation

The most studied hormone related to latex production is ethylene. It had been proposed that ethylene prolongs latex flow and accelerates sucrose metabolism (Tungngoen *et al*., 2009). Ethylene acted on membrane permeability, leading to

prolonged latex flow with the involvement of HbPIP2 (Tungngoen *et al*., 2009). It has been shown that key enzymes involved in rubber biosynthesis are not involved during ethylene stimulation (Zhu and Zhang, 2009). Only the enzymes involved in sucrose metabolism, regulatory enzyme of coagulation, and lutoid stability were involved (Zhu and Zhang, 2009).

2.6 Latex allergy

The healthcare of the employees with latex allergy is a major concern for the industry (Merget *et al*., 2010). Powdered latex gloves or latex gloves with high allergen content have been forbidden in Germany since 1998 (Merget *et al*., 2010). An experiment done on 32 employees showed that 30% of them had latex allergy symptoms (Merget *et al*., 2010).

Daily contact latex products include gloves, rubber-handled racquets, balloons, latex-padded play pits, infant pacifiers, and bottle nipples (Estelle and Simons, 2010). Contact with these products will cause a hypersensitive response to the user (Estelle and Simons, 2010). The proposed mechanism of how the latex allergen triggers the anaphylaxis is the binding of the latex allergen to the immunoglobulin E of the patient (Estelle and Simons, 2010). The identification of latex allergens allows the production of more precise commercial test kits (Lieberman *et al*., 2010).

Natural latex has also been associated as a risk factor of spina bifida (Ozkaya *et al*., 2010). Screening patients for latex allergy prior to surgery is an important but intensive procedure (Khan *et al*., 2010). The appropriate testing strategy for diagnosing latex of *H. brasiliensis* allergy involves *in vitro* specific IgE or skin prick testing (Khan *et al*., 2010). The sensitivity and specificity of both tests are influenced

by patient-specific factors or manufacturing processes that alter the clinically relevant allergens in skin testing solutions (Khan *et al*., 2010).

2.7 Genome sequencing and transcriptome

DNA sequencing data allows a better understanding of the organism in a more complete picture. The genome of *H. brasiliensis* provides scientists with knowledge of the genes that are important for research and development. Transcriptome analysis of *H. brasiliensis* latex has also been performed (Han *et al*., 2000). It showed that the highly expressed genes were rubber elongation factor (REF) and small rubber particle protein (SRPP) (Han *et al*., 2000). The second most frequent transcripts, next to rubber biosynthesis-related genes, were defense genes and protein metabolism-related genes (12.6% each). About 27% of the transcripts were with unknown function (Han *et al*., 2000). A previous experiment of differential gene expression of the leaves and latex showed that each have a distinct gene expression profile. It was found that chitinases, pathogenesis-related protein, phenylalanine ammonia lyase, chalcone synthase, chalcone isomerase, cinnamyl alcohol dehydrogenase, 5-enolpyruvylshikimate-3-phosphate synthase, hydrolytic enzymes, cellulase, and polygalacturonase are highly expressed in the laticifer. Genes for the photosynthesis pathway and chloroplastic form enzymes such as ribulose-bisphosphate carboxylase small subunit and chlorophyll a/b-binding protein are not detected in laticifers. Transcripts for the cytoplasmic form of glutamine synthase are preferentially expressed in laticifers, whereas those for the chloroplastic form of the same enzyme are present mainly in leaves. Other genes such as ATPase, actin, and ubiquitin are equally expressed in laticifers and leaves (Kush *et al*., 1990).

These data were further justified in 1993, when the details of the organelles inside the laticifer cell was published (Jacob *et al*., 1993). Organelles found inside the laticifer cell include rubber particle, Frey Wyssling particles, lutoid, and ribosomes. The nucleus and mitochondria was found in the laticifer but not in latex collected from tapping. Rubber particle is the main factory for the synthesis of natural rubber (cis 1,4-polyisoprene). The content of Frey Wyssling particles include carotenoid and plastochromanol (Whittle *et al*., 1965). Lutoid is the organelle producing lectin-like protein (Gidrol *et al*., 1994; Wititsuwannakul *et al*., 2008a). This lectin-like protein binds to the rubber particle surface protein (Wititsuwannakul *et al*., 2008b), resulting in latex agglutination. The blocking of the transport system by the agglutination stops the out flow of the latex.

2.8 Rubber database creation

The wide commercial applications of natural rubber led to the sequencing of the rubber tree. The rubber tree species sequenced in this project is *Hevea brasiliensis*. The huge amount of data provided from genome sequencing and transcriptomic data prompt for a better visualization of the data.

There are some databases available for public access such as KEGG and AraCyc providing different types of pathways for different organisms. However, a varying number of false-positive predictions occur in these databases. The false positive occurs due to the quality of manual validation and curation of the databases. The lack of consistency in annotation standards and the lack of comparable quality in validation and curation hinder researchers seeking to meaningfully compare the metabolic networks of individual species housed in different databases.

An organism's phenotype is closely dependent on its metabolic activity. Metabolic pathway reconstruction of *H. brasiliensis* focuses on developing the basic metabolic platform with validation from genomic and transcriptome data. The aim is to provide a better visualization of the huge amount of genes predicted from automated gene prediction pipeline and allow more flexible manipulation of the data. With this data manipulation tool and huge amount of data, the underlying mechanism of the cellular process is visualized.

Throughout the years, some small scale attempts had been done to predict and explore the metabolic activity in *H. brasiliensis*. The types of metabolic pathways predicted in small scale vary according to research interest. For example, the occurrence of the cyanogenic glucosides pathways was predicted in 1988, proposing this secondary plant products serve in the metabolism of developing plants as N-storage compounds and do not exclusively exhibit protective functions (Selmar *et al*., 1988). In 2008, the mevalonate pathway synthesizing terpenoids from acetyl-CoA was validated by cloning the enzymes catalyzing this pathway (Sando *et al*., 2008a; Sando *et al*., 2008b). Provided all the information from different research interests, it is hard to get a better view of the metabolic activity in *H. brasiliensis*.

The aims of this project include:

Specific aim 1: Reconstruction of *H. brasiliensis* metabolic pathways using information from *Hevea* genomic and transcriptomic data. The reconstructed metabolic pathways will provide a comprehensive understanding of the metabolic activity at the cellular level.

Specific aim 2: Manual construction and curation of the plant resistance pathways based on published literature. The curated pathways will provide a better view for the molecular interactions at the cellular level. The merging of the curated pathways to the reconstructed metabolic pathways will provide information about the regulation of latex production by plant disease resistance.

Specific aim 3: Network analysis and curation based on the available relations in the rubber database. The network is used to predict the shift in rubber production of *H. brasiliensis* due to application of artificial ethylene to rubber tree.

All the reconstructed pathways, curated pathways and networks were housed in the *H. brasiliensis* single organism specific database.

CHAPTER 3

MATERIALS AND METHODS

3.1 Overview for rubber database creation

Rubber database was created from a set of pathway collections. The pathways were further validated by the genomic and transcriptomic data. The genes of the rubber can be found in the rubber database based on pathway. The pathways reconstructed were further used to predict functional physiological processes in rubber. An overview of the rubber database creation is shown in Figure 3.1.

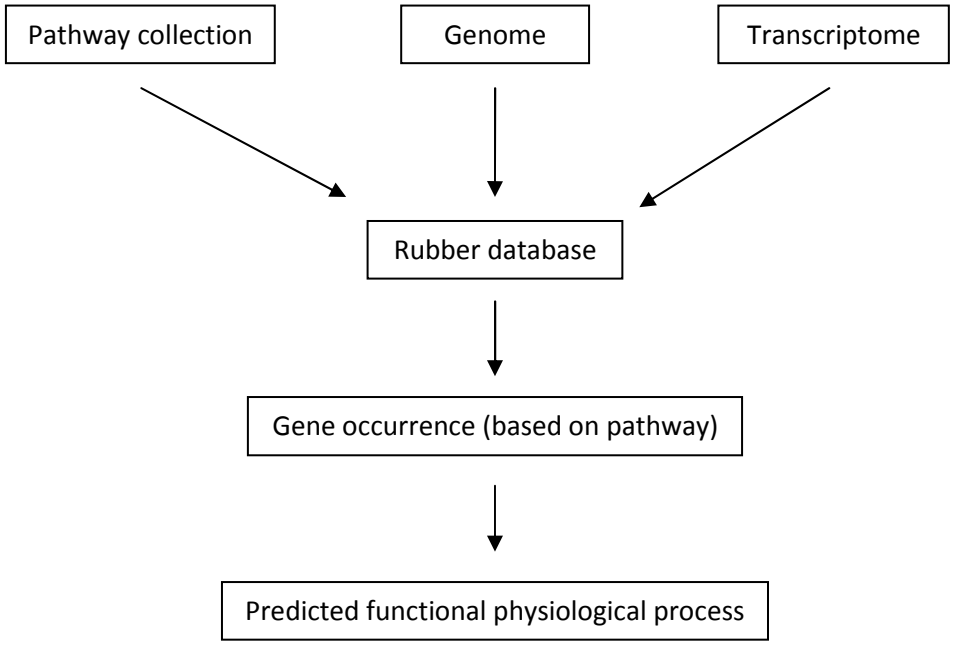

Figure 3.1 Overview of rubber database creation

3.2 Pathway studio implementation process

The metabolic reconstruction of *H. brasiliensis* can be generally divided into three basic components: 1) annotated protein sequences from the genome sequencing pipeline, 2) reference metabolic pathway databases, and 3) mappings of the annotated sequences to pathways in the reference database.

3.2.1 Annotate protein sequences from genome sequencing pipeline

H. brasiliensis machine predicted genes were annotated to the protein sequences. The genes annotated in the rubber database include genes from *Arabidopsis thaliana*, *Ricinus communis*, *Oryza sativa*, and *Populus trichocarpa*.

3.2.2 Reference pathways construction

The reference metabolic and hormonal signaling pathways are Ariadne ResNet plant pathways collection and some newly constructed metabolic pathways based on the publicly accessible database from PoplarCyc. The reference metabolic pathways can be divided into 4 categories. The categories were Biosynthesis, degradation/utilization/assimilation, detoxification, generation of precursor metabolites, and energy and transport pathways. The hormonal signaling pathway is grouped in a separate folder. Figure 3.2 shows the steps involved for a new reference pathway construction. Pathway Studio commands used to construct new reference pathway include:

1. Add> entity – to add a new entity into database
2. Add> New Relation between Selected Entities – to add new relation between selected entities in the database
3. Import> gene list – to import gene list into a pathway

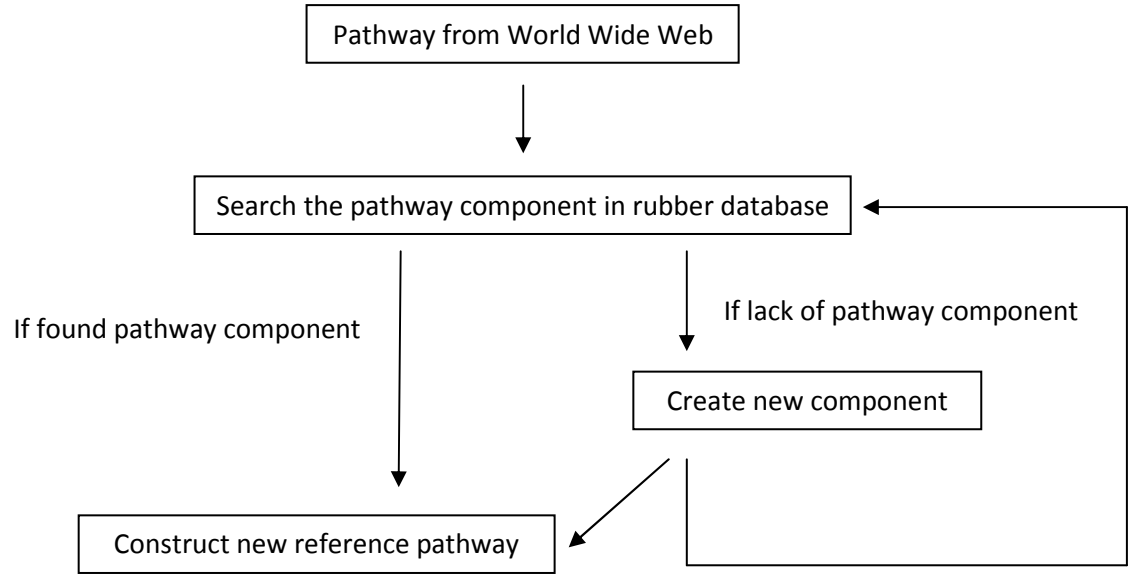

Figure 3.2 Workflow to construct new reference pathway

3.2.3 Mapping of the genes to the reference pathway

The annotated protein sequences from the genome sequencing pipeline is mapped to different ID based on the version of draft genome sequence. The mapping process functions to label different properties to the entities in the rubber database. The different type of properties identifier created based on different versions of the draft genome include 011510 assembly ID, rubber Ortholog ID, rubber ID, assembly 161009 ID, Arabidopsis scaffold assembly ID, Solexa assembly ID, newbler 091109 ID, and EVM 123109 ID. The functional gene set in the leaves is represented as 454 isotig ID. Pathway studio command used to map the data is:

Import> relation from tabular format – to annotate the draft genome and functional gene set

3.3 Manual gap closing of the metabolic pathway

The reference pathways are built based on the other plant metabolic pathways. After the mapping of the genes to the reference pathways, there are still a lot of pathways missing the catalyst for a particular reaction. Manual gap closing of the pathways will ensure that the metabolic pathway occurs in rubber. The gap in metabolic pathway is highlighted in red after the mapping for easy identification. The steps involved in manual gap closing of the metabolic pathway are shown in Figure 3.3. Gene search will be done in the World Wide Web to get the sequence of the appropriate template. The template gene is blast against the rubber draft genome. After the orthologous gene is found in rubber, the gap is closed. If there are other gaps in the pathway, the process is repeated until all the gaps in a particular pathway are closed. Some useful resources in World Wide Web are included in section 3.4. Pathway studio commands used during the pathway reconstruction process include:

1. Edit >copy – copy entities, relations into clipboard.
2. Edit public Object – command in web client to add protein members into a particular functional class (only applicable after rubber database become a server database instead of local database).
3. Edit>paste – paste entities, relations and protein members into a particular functional class.

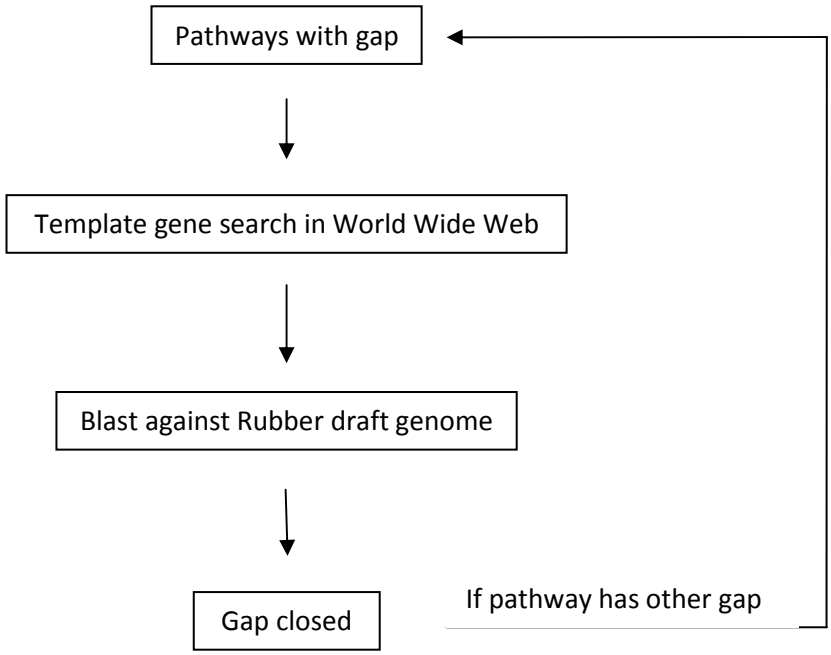

Figure 3.3 Workflow to manually close the gap in a pathway

3.4 Resources

3.4.1 Gene ontology

The Gene Ontology (GO; http://www.geneontology.org) project is a major collaborative bioinformatics initiative that aims to standardize the representation of gene and gene product attributes across species. Gene ontology provides a set of structured, controlled vocabularies for the annotation of genes, gene products, and sequences. The GO ontologies are expanding both in content and in structure throughout the years (Consortium, 2009).

3.4.2 KEGG pathway database

KEGG pathway database (http://www.genome.jp/kegg/pathway.html) is a publicly available collection of manually drawn pathway maps. The pathways available in this database can be divided into metabolism, genetic information processing, environmental information processing, cellular processes, organismal

systems, human diseases, and drug development. The pathways in this database range from unicellular prokaryotes to multicellular eukaryotes. The pathways available are customizable according to organisms of interest (Okuda *et al*., 2008).

3.4.3 National Center for Biotechnology Information (NCBI)

The National Center for Biotechnology Information (NCBI) is one of the world's premier Web sites for biomedical and bioinformatics research. Based within the National Library of Medicine at the National Institutes of Health, USA, the NCBI hosts many databases used by biomedical and research professionals. The services include PubMed, the bibliographic database; GenBank, the nucleotide sequence database; and the BLAST algorithm for sequence comparison, among many others. One of the frequently accessed database is the Entrez Gene.

3.4.3(a) NCBI Entrez Gene

Entrez Gene has been implemented at NCBI to organize information about genes. Each Entrez Gene record is assigned a unique identifier, the Gene ID. Known and predicted genes are established in Entrez Gene record, which are defined by nucleotide sequence or map position. At the Entrez Gene web interface, users are able to visualize detailed information of a particular gene of interest.

3.4.3(b) NCBI PUBChem

PubChem (http://pubchem.ncbi.nlm.nih.gov) is a public repository for biological properties of small molecules hosted by the US National Institutes of Health (NIH). PubChem BioAssay database currently contains biological test results for more than 700,000 compounds. It presents a set of web servers to facilitate and

optimize the utility of biological activity information within PubChem. These web-based services provide tools for rapid data retrieval, integration, and comparison of biological screening results, exploratory structure–activity analysis, and target selectivity examination (Wang *et al*., 2009).

3.4.3(c) NCBI Biosystem

The NCBI BioSystems database (http://www.ncbi.nlm.nih.gov/biosystems/), centralizes and cross-links existing biological systems databases, increasing their utility and target audience by integrating their pathways and systems into NCBI resources. This integration allows users of NCBI's Entrez databases to quickly categorize proteins, genes, and small molecules by metabolic pathway, disease state, or other BioSystem type, without requiring time-consuming inference of biological relationships from the literature or multiple experimental datasets (Geer *et al*., 2010).

3.5 Signaling pathway construction (Plant defense mechanism)

The Plant defense mechanism is constructed by the literature search in World Wide Web. The relations in the pathways are manually curated. The workflow is shown in Figure 3.4.

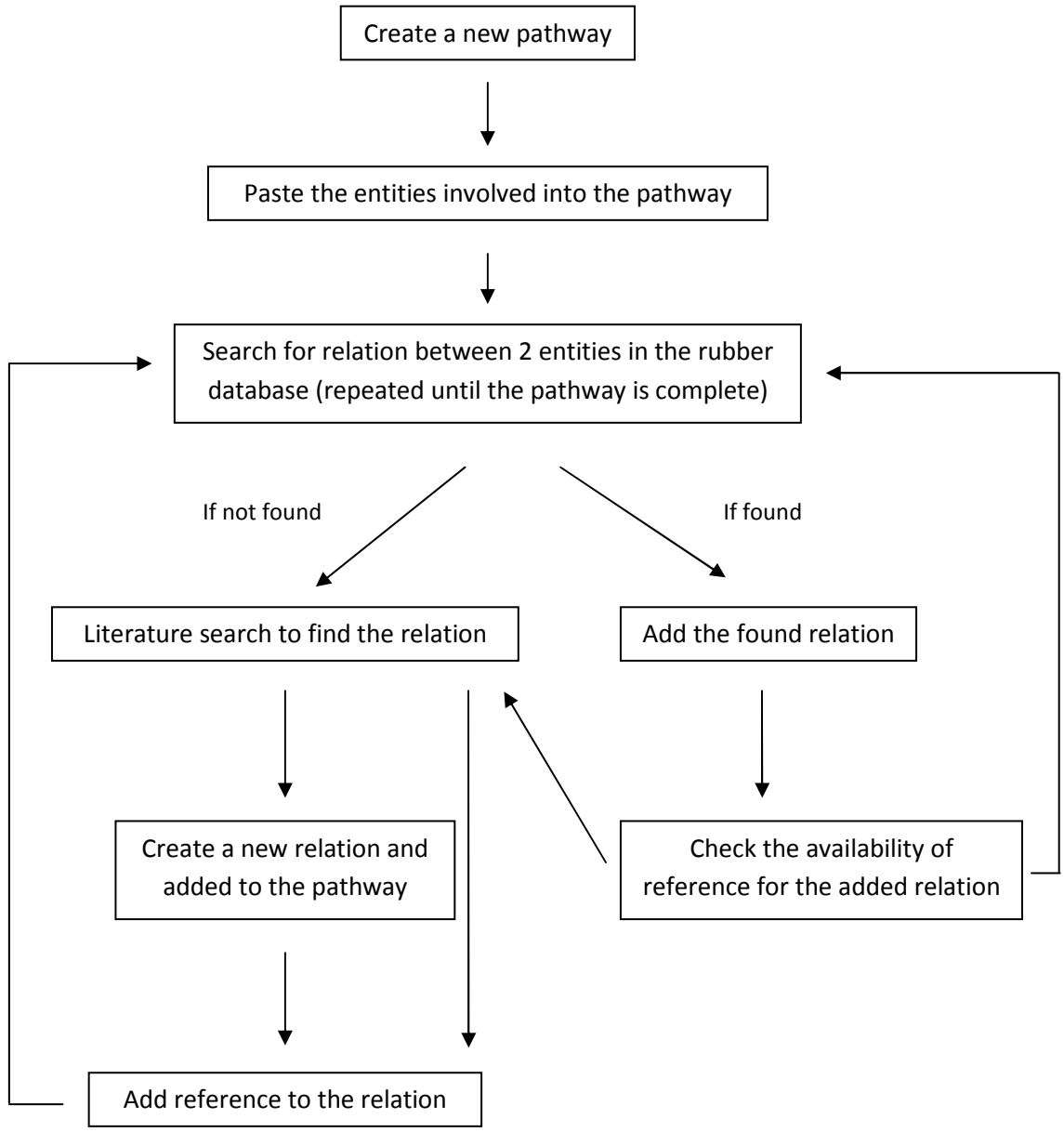

Figure 3.4 Workflow to construct and manually curate a pathway based on publication

Commands involved during the process include:

1. Add> entity – to add new entity that is not found in the rubber database.
2. Add> New Relation Between Selected Entities – to add new relation that is not found in the rubber database.
3. Edit > copy – to copy entity or relation into clipboard.

4. Edit> paste – to paste entity or relation from clipboard into pathway.
5. Properties (of a particular entities/ relation) – to add relevance sentences to support the occurance of a particular relation.
6. View> Relation Table View – To figure out the relation without references.

3.6 Cross-talk between pathways (How necrotic pathogen infection may influence latex production)

Approaches taken to find out possible regulation on the metabolic pathway is the construction of signaling pathways regulating the reconstructed metabolic pathways (Figure 3.5). The relations connecting the entities in the predicted pathway are manually curated.

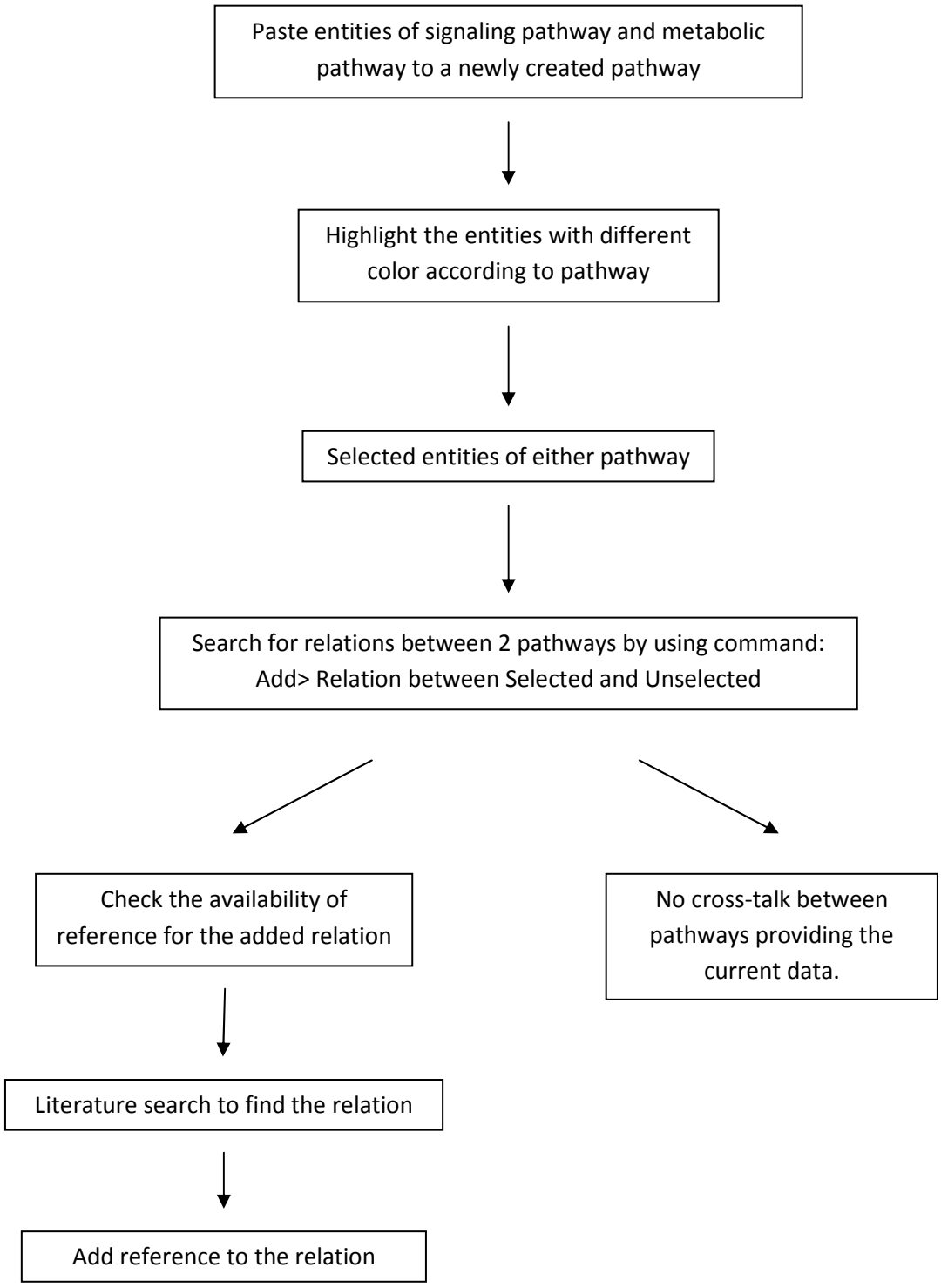

Figure 3.5 Workflow to construct and manually curate cross-talk between 2 pathways

Commands involved during the process include:

1. Add> Relation between Selected and Unselected – to get the appropriate relation applicable to the pathway from the database
2. Edit> copy – to copy a new entity or relation to clipboard
3. Edit> paste – to paste a new entity or relation to a particular pathway
4. Properties (of a particular entities/ relation) – to add relevance sentences to support the occurance of a particular relation.
5. View> Relation Table View – To figure out the relation without references.

3.7 Network analysis (Ethylene regulates sucrose metabolism and Predicted ethylene regulation on water flow into the laticifer)

The other approach taken to more efficiently use the rubber database is the network analysis. In this project, the network analysis is used to figure out other possibility or gap missed out by recent research and provides a better view of the possible paths taken by a single molecule inside *Hevea* cell. By doing so, it allows a better prediction on the function, location, and phenotypic changes caused by the molecule. The workflow is shown in Figure 3.6.

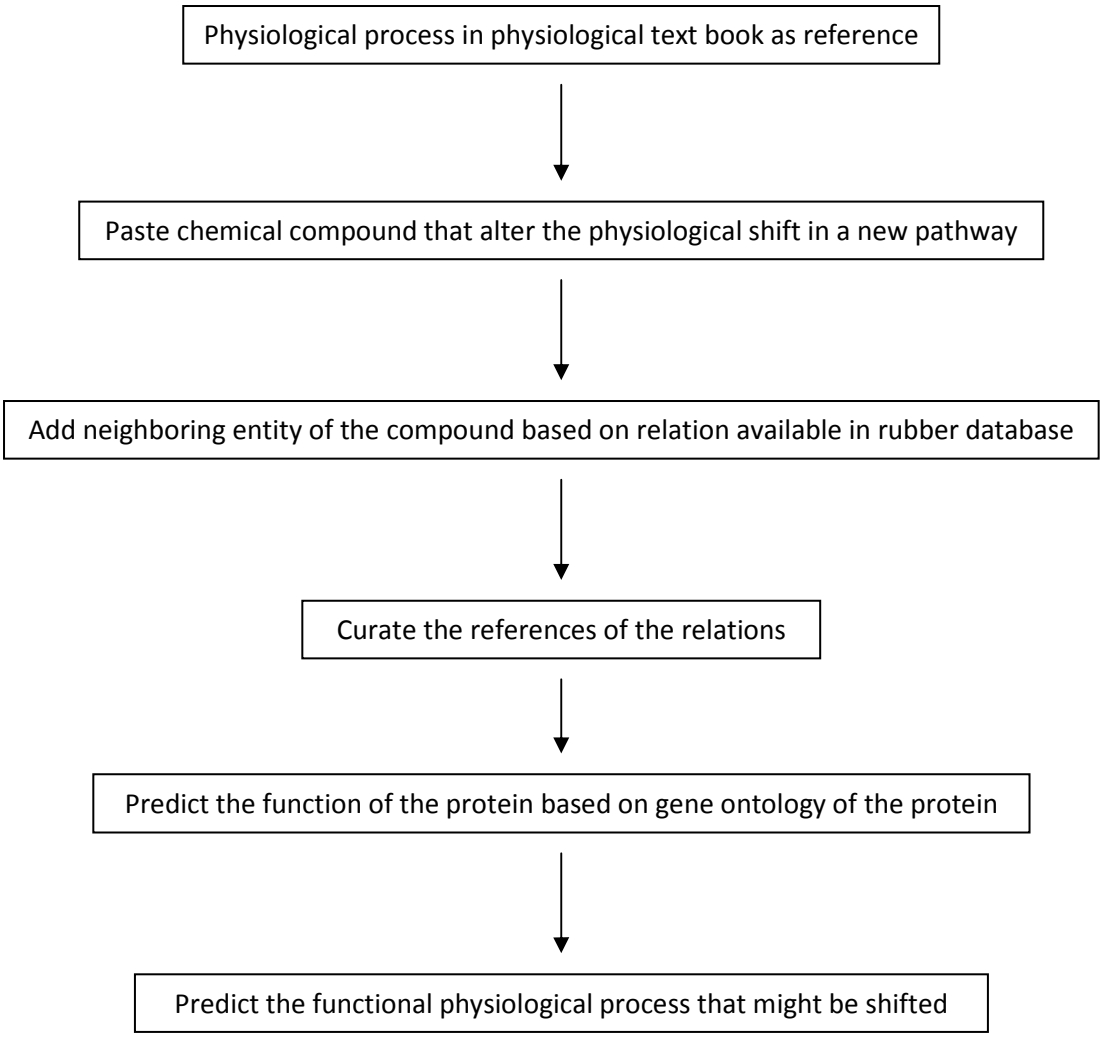

Figure 3.6 Workflow to construct and manually curate a network to predict shift in physiological process

The commands used in this part include:

1. Add> neighbors from DB> All – to get the neighbors entities of the selected entities based on relations available in the database
2. Add> Relation between Selected and Unselected – to get the appropriate relation applicable to a particular pathway from the database
3. Edit> copy – to copy a new entity or relation to clipboard
4. Edit> paste – to paste a new entity or relation to a particular pathway

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Metabolic pathways

4.1.1 Photosynthesis light reaction

One of the pathways generated was the photosynthesis light reaction in Figure 4.1. The photosynthesis light reaction (GO:0019684) takes place in the chloroplast thylakoid membrane (GO:0042651). One of the crucial steps in the photosynthesis light reaction is light harvesting, which involves the absorption and transfer of the energy absorbed from light photons between photosystem reaction centers (GO:0009765). During the photosynthesis light reaction, light energy is harvested and used to power the transfer of electrons among a series of electron donors and acceptors. The final electron acceptor is NADP+, which is reduced to NADPH. Light reactions also generate a proton motive force across the thylakoid membrane, and the proton gradient is used to synthesize ATP. There are two chemical reactions involved in the photosynthesis light reaction: water oxidation in photosystem II, and NADP reduction in photosystem I.

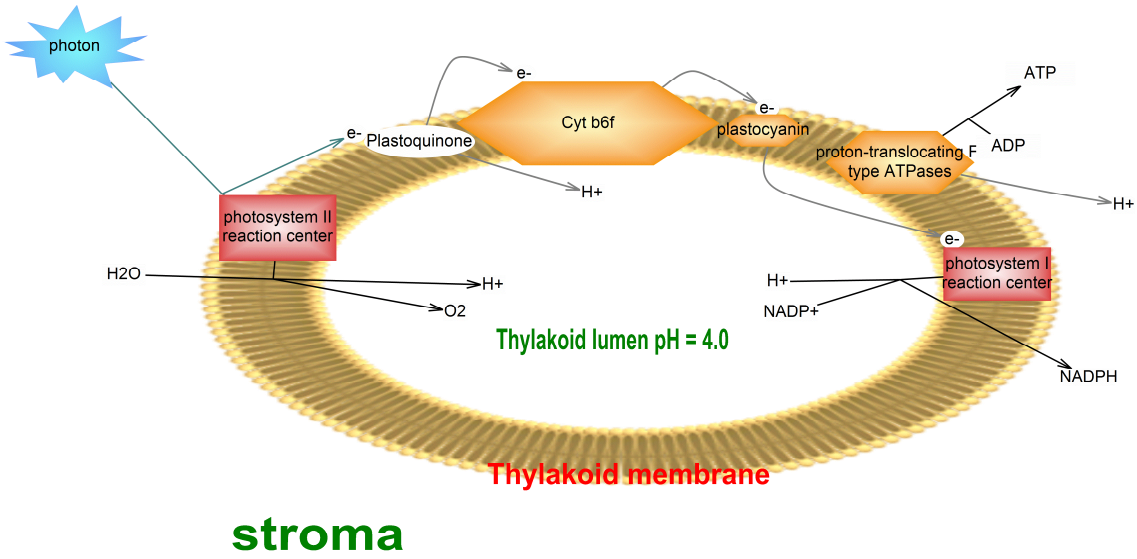

Figure 4.1 Reconstructed photosynthesis light reaction pathway in *H. brasiliensis*

4.1.2 Photorespiration and hydrogen peroxide degradation

Figure 4.2 shows the photorespiration and hydrogen peroxide degradation. Photorespiration (GO:0009853) is defined as a light-dependent catabolic process occurring concomitantly with photosynthesis in plants (especially C3 plants) whereby dioxygen (O_2) is consumed and carbon dioxide (CO_2) is evolved. Photosynthesis will occur when the surrounding CO_2 concentration is high. Photorespiration will occur when O_2 concentration is higher. During the photorespiration process, hydrogen peroxide is synthesized in the peroxisome. This toxic molecule is converted into non-toxic O_2 molecule. This pathway is the results of the merging of the photorespiration in Arabidopsis and superoxide radical degradation of the PoplarCyc pathway. The reconstructed pathway for *Hevea* is the same as reference pathway.

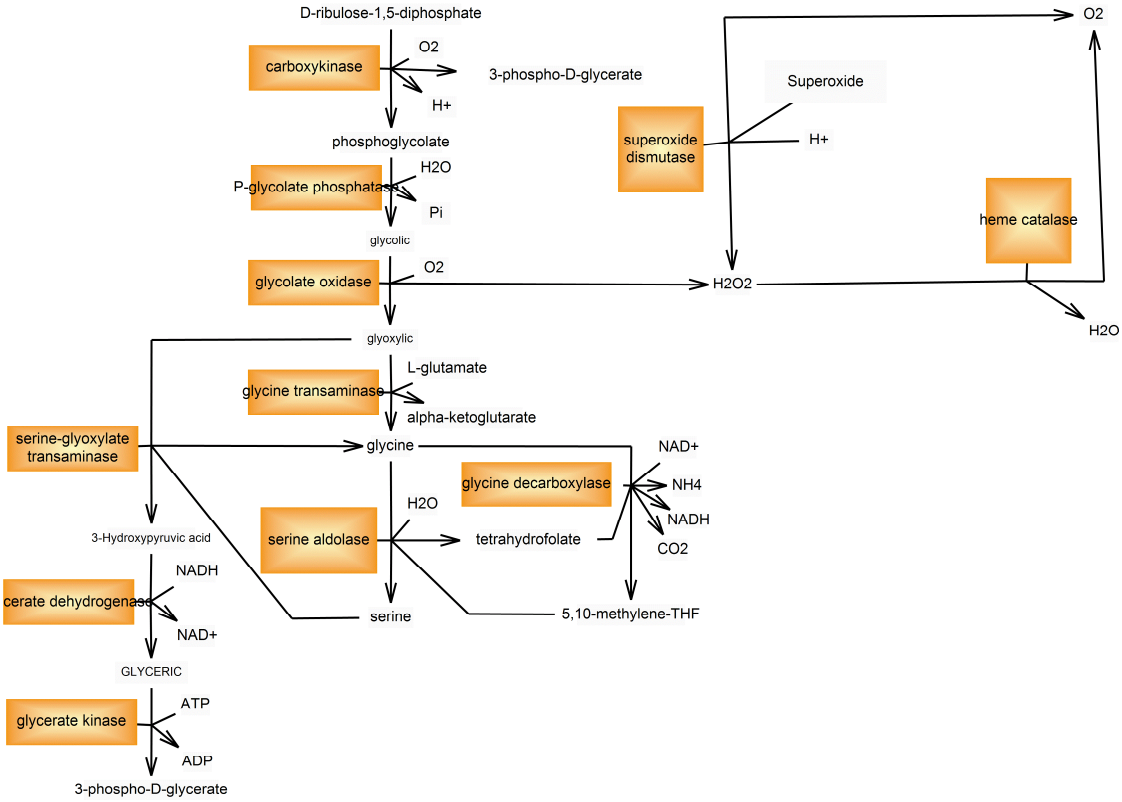

Figure 4.2 Reconstructed photorespiration and hydrogen peroxide degradation pathway in *H. brasiliensis*

4.1.3 Cholesterol biosynthesis

Figure 4.3 illustrates the conversion of farnesyl pyrophosphate to cholesterol. The cholesterol biosynthesis process (GO:0006695) is defined as the chemical reactions and pathways resulting in the formation of cholesterol, cholest-5-en-3 beta-ol, the principal sterol of vertebrates and the precursor of many steroids, including bile acids and steroid hormones. However, validation from *H. brasiliensis* shows that cholesterol biosynthesis occurs in this plant despite being a pathway primarily found in mammalian cells. The cholesterol biosynthesis pathway validated to occur in *H. brasiliensis* is cholesterol biosynthetic process via lathosterol (GO:0033490) where lathosterol is involved as intermediate. The reconstructed pathway for *Hevea* is the same as reference pathway.