# Nanopore Full-length cDNA Sequencing Uncovers Genes Related to High Photosynthesis under High-light Condition in *Hirschfeldia incana*

## MSc Research Practice Report


**Student:**

Run Li

**Student number:**

1053837


**Course code:**

BIS79324

**Administrative Supervisor:**

Prof. M. Eric Schranz

**Daily Supervisor:**

Dr. Nam V. Hoang

**Examiner：**

Prof. M. Eric Schranz


**Institution:**

Biosystematics Group

Wageningen University & Research



**Date:**

2023, December 15

# Acknowledgment

I would like to express my sincere gratitude to my administrative supervisor, Prof. M. Eric Schranz. Prof. Schranz has consistently demonstrated kindness, in-time communication, and genuine consideration for the student's perspective. His support has provided not only academic guidance but also a sense of belonging and spiritual encouragement. I am deeply grateful for the mentorship and understanding that have greatly enriched my learning experience. Your insights and mentorship have been invaluable in shaping my skills as a researcher and helping me to achieve my academic goals.

I would also like to express my sincere gratitude to my daily supervisor, Dr. Nam V. Hoang. Dr. Hoang has imparted invaluable knowledge and methods to me, paying meticulous attention to details and providing numerous constructive suggestions. I am truly appreciative of your guidance and expertise which have significantly contributed to my academic growth and research skills.

In addition, I extend my appreciation to the participants who took part in this study, Patrick Verbaarschot for guiding and supporting me during my lab work, and Ludovico Caracciolo for providing me with the plant materials. I am grateful for your time, effort, and willingness to share your experiences and insights.

# Abstract

This study investigates the molecular responses to long-term high-light (HL) stress in *Hirschfeldia incana*, a Brassicaceae species that exhibits high photosynthetic rates under high irradiance. To do that, we first systematically re-analyzed the previously published Illumina transcriptomic data of the model plant *Arabidopsis thaliana* and *H. incana* treated with HL for both short and long durations. In *A. thaliana*, the long-term HL experiment revealed conserved down-regulation of photosynthetic genes, indicating a common light-stress responsive biology. Notably, certain genes exhibited restoration from down-regulation under prolonged HL exposure, reflecting nuanced adaptive strategies. Robust responses to reactive oxygen species (ROS) were observed, with adaptive fine-tuning over time. Diverse mechanisms in response to drought were also identified, enhancing drought tolerance through sulfate transmembrane transport regulation. For *H. incana*, analogous stress response mechanisms were evident, encompassing drought, salt stress, oxidative stress, and hypoxia. The Illumina dataset exhibits down-regulation of ribosomal RNA (rRNA) activities long-term HL, highlighting the prioritization of cellular functions. Concurrently, up-regulation of amino acid metabolic pathways suggested roles in energy provision or stress-related processes. Employing newly generated Nanopore long-read cDNA sequencing unveiled isoform-level complexities in the transcriptomic data, enhancing our comprehension of HL stress responses in *H. incana*. This research provides valuable insights into plant adaptation under long-term HL, emphasizing both conserved and distinctive regulatory elements across species.



**Keywords:** light stress, plant adaptation, photosynthesis, long-read transcriptome sequencing, *Hirschfeldia incana*.

# Table of Contents

# 1. Introduction

Light is a prevalent environmental factor that significantly impacts photosynthesis. Plants require an optimal amount of light energy for photosynthesis to occur efficiently. However, increasing light intensity and/or duration is not always an effective solution to enhance photosynthetic efficiency. Beyond a certain point, plants reach a state of light saturation, where further increases in light intensity do not lead to a proportional increase in photosynthesis (Garassino et al., 2022; Gitelson et al., 2015; Turner et al., 2003). Instead, excessive light can cause photodamage to the photosynthetic apparatus, leading to reduced efficiency and potentially harming plant health (Apel & Hirt, 2004; Garassino et al., 2022; Huang et al., 2019). Moreover, during the majority of their life cycle, plants commonly experience high light intensity that surpasses their photosynthetic capacity (Mishra et al., 2012). Plants have evolved various mechanisms to adapt and acclimate to light stress, where they exhibit a range of physiological and biochemical responses. Understanding their responses to high-light stress is crucial for developing strategies to improve crop productivity in light-stressed environments.

One of the key strategies employed by plants is photoprotection, which involves dissipating excess energy as heat through processes like non-photochemical quenching (NPQ) (Horton et al., 1994; Takahashi & Badger, 2011). By diverting the surplus energy away from the photosynthetic machinery, plants can minimize the formation of harmful reactive oxygen species and prevent photooxidative damage (Moustakas, 2022; Takahashi & Badger, 2011). Additionally, plants regulate their photosynthetic apparatus through reversible downregulation processes, such as photoinhibition (Aro et al., 1993; Murata et al., 2007; Powles, 1984). Photoinhibition refers to the temporary reduction in the photosynthetic efficiency of plants under high-light conditions. This protective mechanism allows plants to limit the energy absorbed by their photosystems, safeguarding them against potential damage. Furthermore, plants produce a variety of antioxidant compounds, including carotenoids (Havaux, 2014), flavonoids (Fini et al., 2011), and tocopherols (Kobayashi & DellaPenna, 2008), which help neutralize reactive oxygen species generated during excessive light exposure. These antioxidants act as scavengers, preventing the oxidation of important cellular components and maintaining the overall redox balance. Moreover, plants possess specific anatomical features that aid in light regulation. For instance, leaf structures reorientation (Feng et al., 2019; Terashima et al., 2005), stomatal movements (Devireddy et al., 2018), and chloroplast photo-relocation (Maai et al., 2020), which play crucial roles in controlling the amount of incoming radiation, optimizing gas exchange, and protecting photosynthetic apparatus against excess energy. Chloroplast photo-relocation also acts to compensate for protection when NPQ is saturated or absent at extremely high light intensities (Zubik-Duda et al., 2023). While our understanding of the anatomical and physiological aspects of plant responses to high-light stress has significantly advanced, there is still a need to unravel the genetic factors underlying crop resilience and photosynthetic performance in high-light conditions.

Recent studies on the gene expression of model plant species under high-light intensity offer valuable insights into the genetic basis underlying enhanced photosynthetic performance. Huang et al. (2019) employed transcriptomic analysis to investigate the response of *Arabidopsis thaliana* plants subjected to short-term high-light conditions. A large number of genes associated with essential pathways such as photosynthesis, hormonal regulation, and phenylpropanoid metabolism have been identified as highly responsive to high-light stress (Huang et al., 2019). This genome-wide analysis of the light-driven stress responses in the model plant *A. thaliana* not only expands our understanding of the genetic mechanisms underlying plant responses to high-light stress but also paves the way toward interspecies comparative genomic analysis. In this regard, *Hirschfeldia incana*, a $C_3$ species from the

Brassicaceae family, stands out with its remarkable high photosynthetic light-use efficiency even under high irradiance conditions (Garassino et al., 2022) that could be used to compare to pinpoint the underlying genetic factors that contribute to plant light-use efficiency.

*Hirschfeldia incana* is an annual plant species that thrives in waste habitats across the Mediterranean region and extends to Southwest Asia (Siemens, 2010). Its adaptability has also led to its establishment as an adventive species in warm-temperate regions (Siemens, 2010). It serves as an ideal model plant for research, displaying rapid and sustained growth in laboratory conditions (Garassino et al., 2022). Additionally, the *H. incana* genome assembly of approximately 400 Mb in size has been released recently (Garassino et al., 2022), which facilitates genomic analyses and in-depth exploration of its genetic makeup. As an exceptional $C_3$ species that has the highest photosynthetic rates found so far in the Brassicaceae family (Canvin et al., 1980; Taylor et al., 2023), *H. incana* is an exclusively valuable resource for investigating high light-use efficiency and unraveling the associated molecular and genetic basis. However, the mechanisms behind its high photosynthetic performances remain largely unexplored. Most studies on the model plants were associated with short-term responses to high-light treatment (i.e., 72 h), while long-term responses could be important for species adaptation. More recently, by analyzing the expression of a group of selected genes related to photosynthesis in *H. incana* under long-term high-light conditions, Garassino et al. (2022) suggested a potential correlation between the high photosynthetic performance and gene copy number variation (CNV).

Gene copy number variation primarily arises from genomic differentiation events, such as genome duplication, triplication, deletion, or rearrangements (Hastings et al., 2009). Gains or losses of copies can significantly affect the expression of a certain gene, resulting in associated phenotypic variation and adaptation to different environmental conditions, which drives evolutionary divergence and facilitates functional specialization (Hastings et al., 2009; Town et al., 2006; Żmieńko et al., 2014). *Hirschfeldia incana*, together with *Brassica* species, underwent a whole-genome triplication event after their divergence from *A. thaliana*, which led to changes in their genome structures (i.e., three subgenomes) and gene expression (i.e., biased expression of duplicated gene copies from the three subgenomes) (He et al., 2021; Lysak et al., 2005; Perumal et al., 2020; Schranz et al., 2006; Wang et al., 2011). By comparing the conserved 260 photosynthesis-related genes between *H. incana* and *A. thaliana*, Garassino et al. (2022) found that the increased copy numbers of photosynthesis- and photoprotection-associated genes in *H. incana* relative to *A. thaliana* are positively correlated to their high photosynthetic efficiency. Besides, the same correlation was also observed between *A. thaliana* and two close relatives of *H. incana*, namely *Brassica rapa* and *B. nigra* (Garassino et al., 2022). Furthermore, most of the selected genes exhibited significantly higher expression levels under high-light condition (1,800 μmol $m^{-2}$ $sec^{-1}$) compared to growth condition (200 μmol $m^{-2}$ $sec^{-1}$) when copy number increased. While no consistently higher gene expression levels were observed in either *B. rapa*, *B. nigra*, or *H. incana* compared to *A. thaliana* when testing additional photosynthesis-related genes that are present as single copy across all four species, which shed light on the necessary role of gene copy number extension in the high photosynthetic performance (Garassino et al., 2022). However, previous studies have predominantly relied on short-read sequencing, leading to inadequate investigation due to their inability to distinguish isoforms derived from alternative splicing (AS). This limitation is particularly pronounced in plant genomes marked by frequent whole-genome duplications, introducing complexities in read assignment (Bernard et al., 2014; Mehrotra & Goyal, 2014). In plants, AS plays a pivotal role in regulating plant development and stress responses (Filichkin et al., 2018). As exemplified, isoforms within the plant GPX (*GLUTATHIONE PEROXIDASE*) gene family demonstrate varied tissue-specific expression patterns and responses to environmental stressors (Bela et al., 2015). Hence, it remains uncertain whether

high-light stress would induce differentially expressed isoforms which could potentially contribute to their exclusive role in high photosynthetic performance of *H. incana*. Besides, the physiological and genetic basis of more efficient photosynthesis under long-term high-light conditions remains largely unexplored, highlighting the need for further investigation to elucidate the mechanisms underlying the plant's adaptation to high-light stress. Therefore, the objective of this study was to investigate the adaptation mechanism of plants to long-term high-light at the isoform-level. This was done through a Nanopore long-read transcriptome analysis of *H. incana* plants grown under long-term low vs. high-light conditions, as well as a comparative analysis of differential gene expression in *A. thaliana* under long-term and short-term high-light conditions. I hypothesized that long-term high-light stress induces differential expression of photosynthesis-related gene copies and/or isoforms.

# 2. Materials and Methods

## 2.1. Plant Materials and Short-read Illumina Data Acquisition

The plant materials for this study consisted of *H. incana* accession 'Nijmegen.' Briefly, plants from two treatment groups (high-light and low-light) were used. The high-light (HL) group consisted of plants exposed to a continuous light intensity of 1,800 µmol $m^{-2}$ $s^{-1}$ for six weeks, while the low-light (LL) group comprised the same growth stage plants grown under a reduced light intensity of 200 µmol $m^{-2}$ $s^{-1}$. Six biological replicates were included for each treatment group, and leaf samples were collected from the 3rd-5th compound leaves counting from the top of each plant (referred to as functional leaves). During leaf sample collection, each selected leaf was carefully detached from the plant using sterile forceps. The samples were immediately placed in pre-labeled, sterilized centrifuge tubes to prevent contamination. The tubes were sealed, snap-frozen in liquid $N_2$ and stored at -80 °C until further processing.

Existing data sets were utilized, including *A. thaliana* short-term HL gene expression data (Huang et al., 2019) (hereinafter short-term experiment), *A. thaliana* and *H. incana* long-term HL gene expression data (Garassino et al., in prep). The *A. thaliana* short-term and long-term gene expression data were analyzed to compare the plant responses under different durations of high-irradiance exposure. The *H. incana* long-term gene expression data represented the gene expression of the whole leaf canopy, which was compared with the new sequencing data generated from functionally developing leaves.

## 2.2. RNA Isolation and cDNA Library Preparation for Long-read Nanopore Sequencing

RNA isolation was performed following the TRIzol™ Reagent RNA isolation protocol combined with the RQ1 RNase-Free DNase Protocol. Approximately 100 mg of each sample was ground using a mortar and a pestle. The quality and quantity of the extracted RNA were assessed using a NanoDrop spectrophotometer and were evaluated by gel electrophoresis. RNA quality evaluation was based on the presence of intact ribosomal RNA (rRNA) bands (i.e., 28S and 18S), absence of degradation products (smears or smudging), and overall band intensity. High-quality RNA samples were characterized by well-defined bands. Ultimately, RNA samples with high purity (A260/A280 ratio ≥ 1.8) and qualified gel electrophoresis results were selected for library preparation.

The Nanopore sequencing library preparation was conducted following the PCR-cDNA Barcoding Kit (SQK-PCB111.24) protocol. Firstly, the extracted RNA underwent reverse transcription to synthesize complementary DNA (cDNA). Then, PCR barcoding was carried out to increase the amount of the library material and to incorporate unique barcodes into the cDNA library fragments. PCR conditions were optimized to ensure efficient and unbiased amplification (**Table 1**). After barcoding, adapter addition was performed on the cDNA library, where a unique barcode sequence was attached to each library fragment. The final adapter-ligated libraries underwent purification using magnetic bead-based methods to remove unincorporated primers and undesired fragments. Sequencing was performed on the MinION Mk1B platform using flow cells (FLO-MIN106D) and was base-called with Dorado simplex (https://github.com/nanoporetech/dorado).

**Table 1. The barcoding PCR procedure and conditions.**

| Cycle step | Temperature | Time | No. of cycles |
|---|---|---|---|
| Initial denaturation | 95°C | 30 sec | 1 |
| Denaturation | 95°C | 15 sec | 16 |
| Annealing | 62°C | 15 sec | 16 |
| Extension | 65°C | 4 min | 16 |
| Final extension | 65°C | 6 min | 1 |
| Hold | 4°C | ∞ | |

## 2.3. Reads Alignment and Transcript Quantification

Chopper (https://github.com/wdecoster/chopper) was used for trimming the raw reads, which were then mapped to the reference genome and transcriptome of *H. incana* (https://www.bioinformatics.nl/hirschfeldia/) using Minimap2 v2.17 (Li, 2018). Transcript quantification was performed using NanoCount (Gleeson et al., 2021) and Bambu (Chen et al., 2023), gene expression levels were calculated using transcripts per million (TPM) and counts per million (CPM) values, respectively.

## 2.4. Identification of Differentially Expressed Genes (DEGs)

The Bioconductor package DESeq2 (Love et al., 2014) in the R programming language was utilized for DEG identification. DEGs were determined based on the adjusted *p*-value and expression fold-change. Specifically, DEGs with an adjusted *p*-value smaller than 0.05 while exhibiting an absolute fold change greater than or equal to 2 and/or 1.5 were considered significantly differentially expressed.

## 2.5. BLAST gene search and Analysis of Enriched Gene Ontology (GO) Terms

The *A. thaliana* gene homologs that mapped to the total *H. incana* genes were used as a background for analyzing the GO enrichment terms of *H. incana* DEGs. The *A. thaliana* homologs of *H. incana* genes were identified by NCBI BLAST on Galaxy (Cock et al., 2015). GO terms and co-expression networks were generated by using ShinyGo (https://bioinformatics.sdstate.edu/go/) and STRING database v12 (Szklarczyk et al., 2020).

# 3. Results

In this study, I conducted a comparative transcriptomic analysis to compare the adaptive responses of *A. thaliana* to short-term and long-term HL stress. Furthermore, I utilized Illumina data from HL-treated canopy samples to investigate the robust photosynthetic capabilities of *H. incana*. To gain deeper insights into HL-induced responses at the isoform level, I performed Nanopore long-read transcriptome analysis on leaf samples of *H. incana*.

## 3.1. Comparative Transcriptomics of Short- and Long-term Light-stress Responses in the Model Plant *A. thaliana*

The short-term HL experiment set the LL and HL conditions at 60 mmol $m^{-2}$ $s^{-1}$ and 1,200 mmol $m^{-2}$ $s^{-1}$, respectively (Huang et al., 2019). Gene expression levels were measured in seedlings harvested at six different time points ranging from 0.5h to 72h, resulting in the identification of a total of 5,403 differentially expressed genes (DEGs) (|fold change| $\geq$ 2, adjusted *p*-value < 0.05) in at least one time point. Accordingly, after exposure to HL stress, plants exhibit activation of stress response pathways such as abscisic acid (ABA) and jasmonic acid (JA) biosynthesis, concurrently inducing anthocyanin biogenesis (Huang et al., 2019). Conversely, repression is observed in nucleotide metabolic processes, brassinosteroid (BR) biosynthetic pathways, as well as genes associated with photosynthesis and growth (Huang et al., 2019). Subsequently, I compared the Arabidopsis short-term response to light stress reported in the above paper with its long-term response (Garassino et al., in prep) to delineate adaptation-related responses from acclimation-related ones.

The long-term HL experiment set 200 mmol $m^{-2}$ $s^{-1}$ as LL condition, and the HL intensity was elevated to 1,800 mmol $m^{-2}$ $s^{-1}$ with an extended exposure duration of six weeks. The total number of DEGs identified was 3,328 and 1,727 for a |fold change| of 1.5 and 2 (adjusted *p*-value < 0.05), respectively (**Figure 1A**). To ensure a more comprehensive understanding of the biological responses, the DEGs with a |fold change| of 1.5 were selected for further downstream analysis. Enriched GO terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were then generated by ShinyGo. The top 20 highly significant GO terms were determined (**Figure 1B**). The up-regulated DEGs were found to be significantly enriched in *alpha*-linolenic acid metabolism and phenylpropanoid biosynthesis, according to the KEGG pathway analysis. Moreover, the enriched biological processes included sulfate transport, response to reactive oxygen species (ROS), response to hypoxia, response to fatty acid, and response to heat. On the other hand, for the down-regulated DEGs, KEGG pathway analysis revealed significant enrichment in photosynthesis, glucosinolate biosynthesis, plant-pathogen interaction, and plant hormone signal transduction. Additionally, the enriched biological processes encompassed light harvesting in photosystem I, response to radiation, growth, and auxin signaling.

Furthermore, I conducted a screening step on the short-term gene expression profile to identify DEGs that exhibited consistent up- or down-regulation from 12h to 72h. A total of 2,146 DEGs were identified, out of which 1,593 genes show overlap with long-term experimental genes (**Figure 1C**). The overlapped genes consist of 711 consistently regulated DEGs, 72 DEGs with reversed regulation patterns, and 810 genes that show restored expression levels (**Figure 1C** and **D**). In contrast, the resemblances and disparities can predominantly be encompassed within five biological processes including photosynthetic processes, light-stress response, ROS response, drought response, and hormonal signaling pathways. To be specific, both long-term and short-term HL caused a decrease in the expression of photosynthetic genes, whereas a lower extent was observed under long-term HL with several photosynthesis-related genes restored from down-regulation (**Figure 1E**). Similarly, the down-regulated response to radiation, light

intensity, and light stimulus in short-term HL conditions remained unaltered after prolonged exposure to HL. In response to ROS, the plants demonstrated robust reactions in both short-term and long-term HL conditions. However, a few up-regulated oxidative stress genes were restored under prolonged exposure to HL. In response to drought, plants with short-term HL and long-term HL exposure exhibited distinct mechanisms. Under prolonged exposure to HL, drought tolerance is enhanced through the regulation of sulfate transmembrane transport, while the number of up-regulated salt stress-related DEGs decreased (**Figure 1B** and **E**). Conversely, short-term HL induces more immediate responses to water deprivation and salt stress (**Figure 1E**). Here, in response to hormonal stimuli, both short-term and long-term HL induced the biosynthesis of zeatin while suppressing growth-related auxin signaling and transport. However, prolonged exposure did not significantly up-regulate the response to ABA and JA but induced an up-regulation in molecular function associated with ABA binding. Moreover, I also observed other intriguing biological changes. For instance, the up-regulation of flavonoid biosynthesis and metabolism in response to short-term HL exposure was not evident under long-term HL conditions, while the isoprenoid activity was up-regulated.

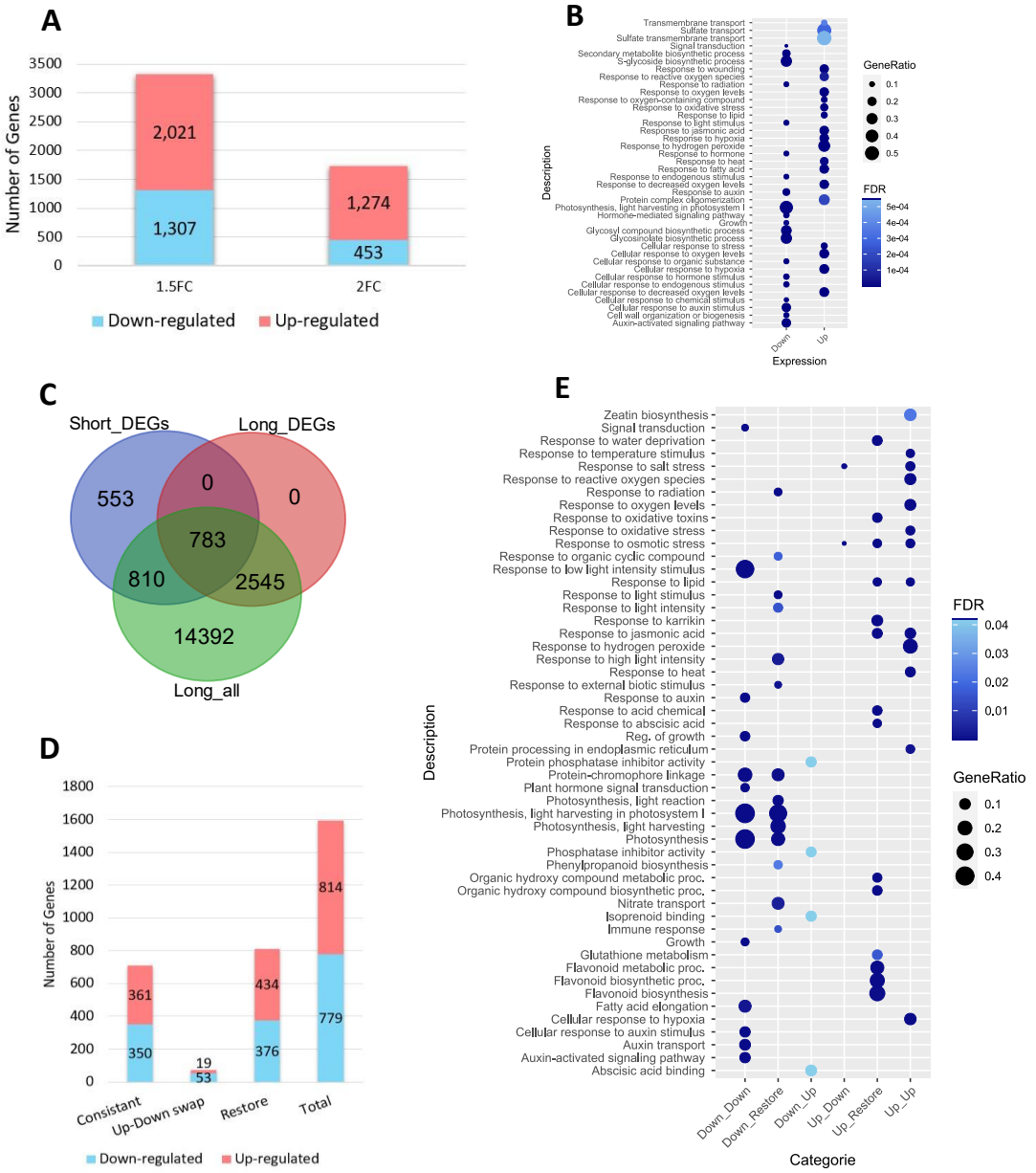

***Figure 1.*** *Differentially expressed genes (DEGs) analysis and comparison between short-term and long-term high-light (HL) conditions.* ***A,*** *the number of 1.5- and 2-fold change (FC) DEGs under long-term HL;* ***B,*** *up- and down-regulated top 20 long-term high-light induced GO enrichment items, determined by gene ratio (number of DEGs divided by the total number of genes in the pathway) and adjusted false discovery rate (FDR) values;* ***C,*** *Venn diagram shows the overlap between short-term HL DEGs and long-term HL genes;* ***D,*** *categories of overlapped genes between short-term and long-term HL experiments. 'Consistent', DEGs that were consistently up- or down-regulated in both conditions. 'Up-Down swap', DEGs exhibiting opposite regulatory patterns. 'Restore', short-term HL-induced DEGs remain unaltered under long-term HL condition;* ***E,*** *the dot-plot displays the GO enrichment items of the long-term HL genes overlapped with the short-term DEGs. 'Down-Down', down-regulated genes under both short-term and long-term HL. 'Down-Restore', unaltered long-term HL genes while being down-regulated under short-term HL. 'Down-up', up-regulated long-term HL genes while being down-regulated under short-term HL. 'Up-Down', down-regulated long-term HL genes while being up-regulated under short-term HL. 'Up-Restore', unaltered long-term HL genes while being Up-regulated under short-term HL. 'Up-Up', down-regulated genes under both short-term and long-term HL.*

## 3.2. Illumina Short-read Sequencing Data Provided Insights into The Underlying Mechanisms Responsible for The High Photosynthetic Capacity in *H. incana*

The subsequent analysis involved examining an Illumina sequencing dataset of *H. incana* whole-canopy RNA samples subjected to long-term HL, encompassing expression data for 23,899 genes. With a threshold of 1.5-fold change, 6,108 DEGs were identified, comprising 2,899 up-regulated and 3,209 down-regulated genes (**Figure 2A**). To investigate the biological mechanisms using the available knowledge, I utilized a BLAST analysis to map the genes of *H. incana* onto their corresponding homologs in *A. thaliana*. In total, 15,474 *A. thaliana* homologs mapped to 23,534 *H. incana* genes, in which 4,083 *A. thaliana* genes mapped to 6,019 *H. incana* DEGs identified above. Specifically, 2,471 *A. thaliana* genes exhibited homology with 2,846 up-regulated *H. incana* genes while another set of 2,449 *A. thaliana* genes showed homology with 3,173 down-regulated *H. incana* genes (**Figure 2A**). Downstream analysis of the DEGs revealed significant regulation of various biological processes in *H. incana* responding to long-term HL stress. Among the down-regulated genes, the top 20 GO enrichment items primarily pertained to ribosome and rRNA activities (**Figure 2B**). Additionally, translation, DNA replication, porphyrin, and chlorophyll metabolism were also found to be down-regulated. On the other hand, the up-regulated genes exhibited a notable enrichment in stress response pathways such as water and salt stress, heat shock, hypoxia, and bacterium-induced responses. Furthermore, several amino acid metabolic pathways along with ABA-mediated responses, glutathione metabolism, pantothenate biosynthesis, and CoA biosynthesis were also observed to be up-regulated. Subsequently, utilizing the *A. thaliana* homologs, I conducted an analysis of the co-expression network among the *H. incana* DEGs within the pathways associated with light stress, water deprivation, and growth. The top 10 most connected DEGs within each pathway were identified based on their node degree (i.e., the number of connections to other genes) to determine their importance in these biological processes (**Figure 2C-E**). In general, the majority of the top DEGs in the three pathways exhibited CNVs, in which differentially expressed copies were observed (**Figure 2F**). Specifically, under long-term HL conditions, the light-stress responsive pathway showed up-regulated enrichment, whereas down-regulation was observed on the *H. incana* copies of *CAM7* (*CALMODULIN 7*, AT3G43810), *MYC2/JAI1* (*JASMONATE INSENSITIVE 1*, AT1G32640), *PAL1* (*PHE AMMONIA LYASE 1*, AT2G37040), and *PAL2* (*PHE AMMONIA LYASE 2*, AT3G53260). Similarly, the water deprivation responsive pathway of *H. incana* also exhibited up-regulation under long-term HL, while copies of *P5CSA* (*DELTA-1-PYRROLINE-5-CARBOXYLATE SYNTHASE A*, AT2G39800), *ABI1* (*ABA INSENSITIVE 1*, AT4G26080), and *SLAC1* (*SLOW ANION CHANNEL-ASSOCIATED 1*, AT1G12480) showed down-regulation.

For plant growth, the topmost connected 10 DEGs include regulators of leaf, floral, and seed development. The majority of these DEGs were found to be down-regulated under long-term HL condition; however, it is noteworthy that copies of *PHYA* (*PHYTOCHROME A*, AT1G09570) and *CYCO3-1* (*CYCLIN-D3-1*, AT4G34160) exhibited up-regulation. Additionally, *AP2* (*FLORAL HOMEOTIC PROTEIN APETALA 2*, AT4G36920), known as a regulator for sepal and petal development, displayed up-regulation in all identified *H. incana* copies.

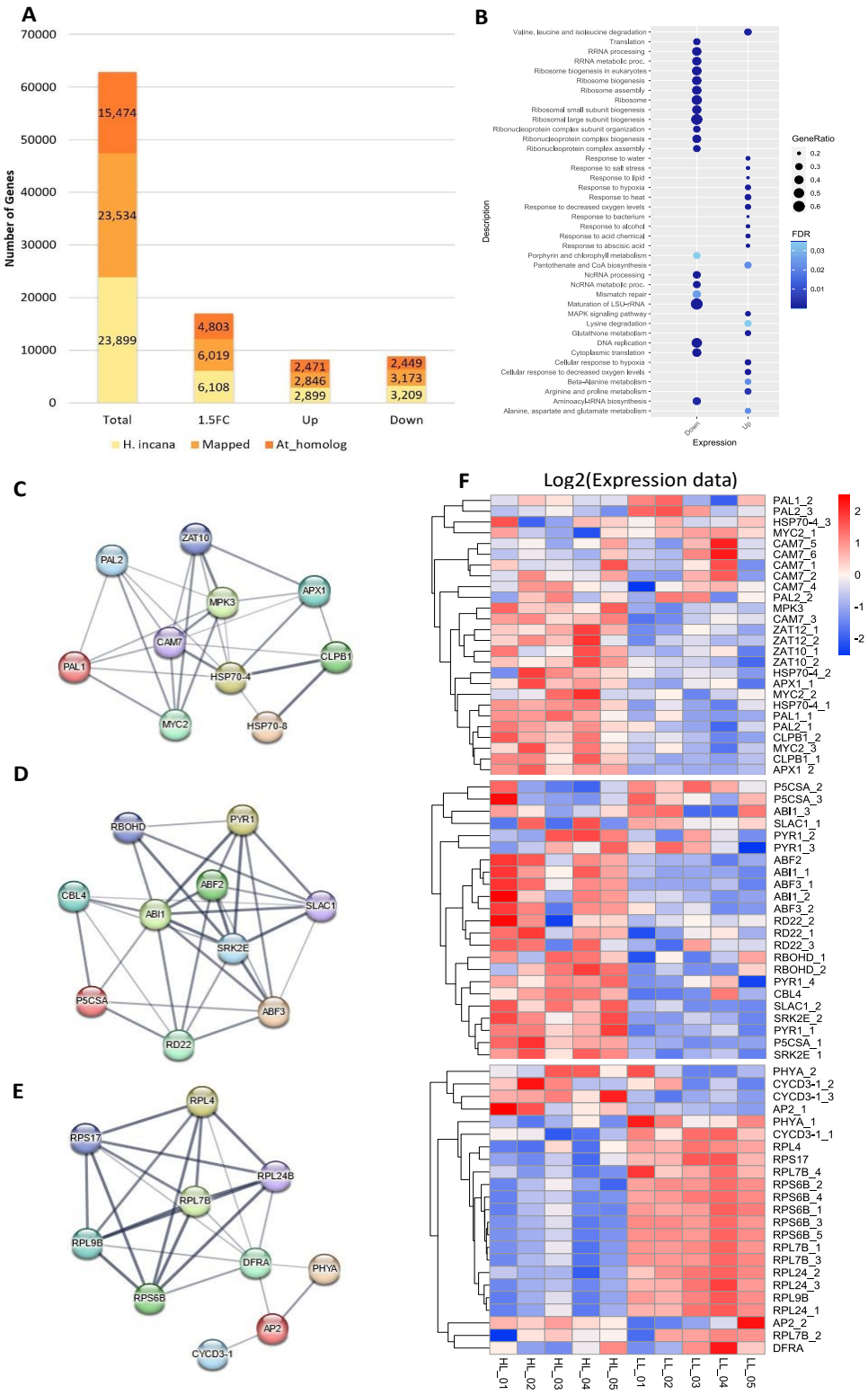

***Figure 2.*** *Analysis of H. incana Illumina dataset.* ***A,*** *gene expression profiling. '1.5FC', long-term HL-induced H. incana DEGs with 1.5 expression fold changes; 'Up', 1.5FC up-regulated H. incana genes; 'Down', 1.5FC down-regulated H.incana genes; 'Mapped', number of H. incana genes having A. thaliana homologs; 'At_homolog', number of A. thaliana homologs obtained by BLAST.* ***B,*** *up- and down-regulated top 20 long-term high-light induced GO enrichment items, determined by gene ratio (the number of DEGs divided by the total number of genes in the pathway) and adjusted false discovery rate (FDR) values;* **C,** *the co-expression network displaying the A. thaliana homologs of the top 10 H. incana DEGs in the light-stress responsive pathway;* **D,** *the co-expression network displaying the A. thaliana homologs of the top 10 H. incana DEGs in water deprivation responsive pathway;* **E,** *the co-expression network displaying the A. thaliana homologs of the top 10 growth-related H. incana DEGs;* **F,** *the heatmap displaying an expression of the top 10 H. incana DEGs' copies in light-stress, water deprivation and growth-related pathways. The heatmap is clustering the log2(expression data) of each gene. 'HL_01-05', five HL biological replications. 'LL_01-05', five LL biological replications. '_number' represents copies of a specific gene.*

## 3.3. The Generation of Nanopore Full-length cDNA Sequencing data set for *H. incana*

To obtain full-length transcripts, I performed Nanopore long-read cDNA sequencing. In total, ~12 Gb cDNA data were generated from the six samples. The datasets were processed to determine several key statistics, which include the read quality, the number of reads, and read length (**Table 2, Figure 3**). Generally, the number of reads varied substantially between the samples. Among them, sample LL_02 contained the highest number of reads, exceeding 4 million reads, while other samples had relatively lower read counts, with HL_03 having the fewest reads at around 2.7 million. In terms of the read quality and length, the statistics were relatively consistent across the samples, with the mean quality scores around 11 and read length predominantly distributed within the ranges of 300 base pairs (bp) and 1,000bp, which is consistent with the sequencing library electrophoresis gel (**Table 2, Figure 3A** and **D**). Additionally, half of the total assembly length was represented by contigs or scaffolds of 800bp or longer as indicated by their N50 values (**Figure 3C**).

**Table 2. Nanopore full-length cDNA-sequencing read data statistics.**

| | HL_01 | HL_02 | HL_03 | LL_01 | LL_02 | LL_03 |
|---|---|---|---|---|---|---|
| Mean read length (bp) | 620.3 | 661 | 627.5 | 674.1 | 666.4 | 740 |
| Mean read quality | 11.2 | 11.3 | 11.1 | 11.2 | 11.1 | 11.2 |
| Median read length (bp) | 488 | 576 | 507 | 617 | 602 | 730 |
| Median read quality (bp) | 12.3 | 12.4 | 12.2 | 12.3 | 12.2 | 12.3 |
| Number of reads | 3,639,033 | 3,020,894 | 2,728,848 | 3,393,871 | 4,106,293 | 2,731,762 |
| Read length N50 (bp) | 815 | 845 | 814 | 872 | 855 | 926 |
| Total bases | 2,257,200,154 | 1,996,880,119 | 1,712,484,957 | 2,287,724,671 | 2,736,296,053 | 2,021,381,407 |

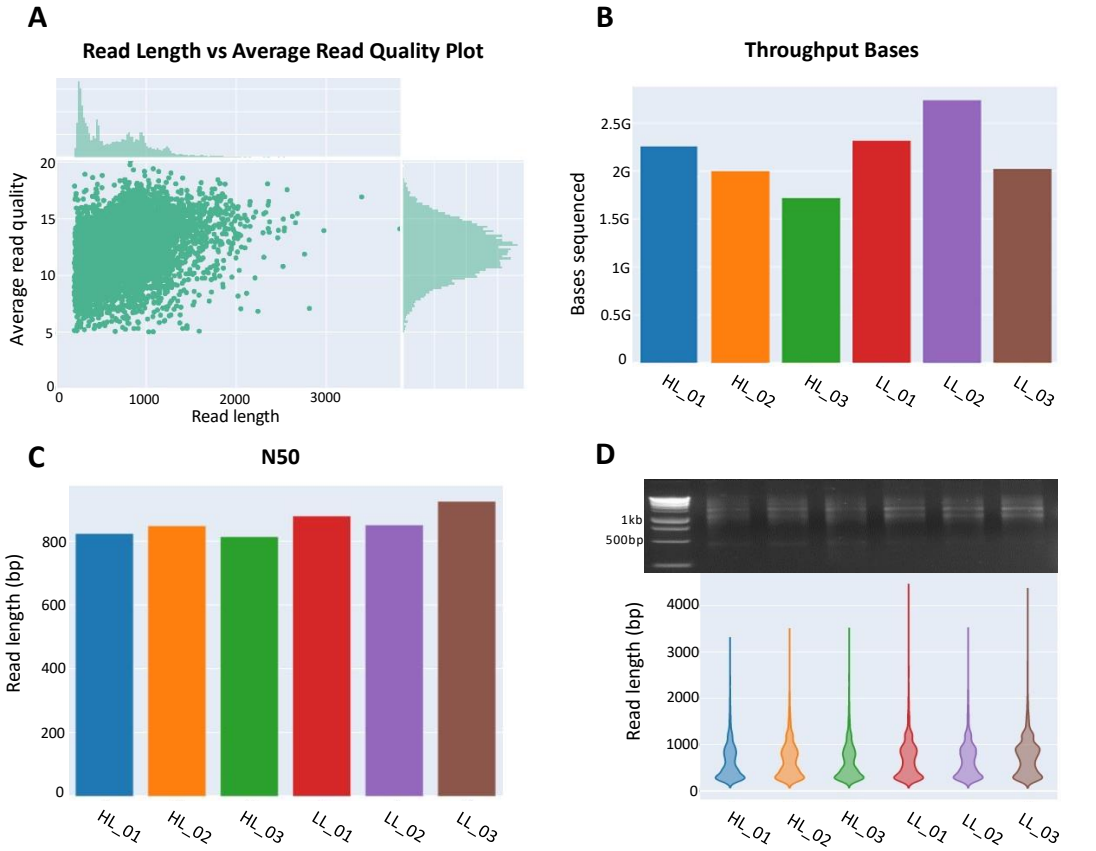

***Figure 3.*** *Nanopore RNA sequencing statistics.* ***A,*** *general read length and quality distribution dot plot;* ***B,*** *comparison of sequenced sample size in bases bar-chart;* ***C****, comparison of N50 values across samples bar-chart;* ***D,*** *comparison of sequencing libraries and read length distribution across samples. A 1kb DNA ladder was used. HL and LL denote high-light and low-light, respectively.*

## 3.4. Transcript Profiling Based on the Full-length cDNA Sequencing Data

In this study, I employed two long-read analysis pipelines, 'NanoCount' and 'Bambu', to quantify the gene expression. The resulting expression levels, categorized into six intervals (>100, 50 to 100, 10 to 50, 5 to 10, 1 to 5, and 0 (excluded) to 1), are presented for each sample analyzed. It is noteworthy that 'NanoCount' employed TPM for expression normalization, while 'Bambu' utilized CPM. Both normalization strategies ensured that the expression values were comparable across samples, accounting for differences in library size and allowing for meaningful cross-sample comparisons. Altogether, 22,569 and 24,347 expressed genes were detected by 'NanoCount' and 'Bambu', respectively. The two pipelines exhibit similar expression patterns with slight differences. To elaborate, according to the quantification output from 'NanoCount', the LL samples showed a consistent distribution across expression intervals, with prominent abundance in the 1 to 10 intervals. Similarly, the LL sample expression from 'Bambu' was also concentrated in the 1 to 10 intervals, whereas an elevated number of genes presented in both tails (0 to 1, >100) of the intervals (**Figure 4**). In terms of HL samples, the 'NanoCount' and 'Bambu' samples exhibited a similar expression profile, with a substantial number of genes in the 5 to 50 intervals.

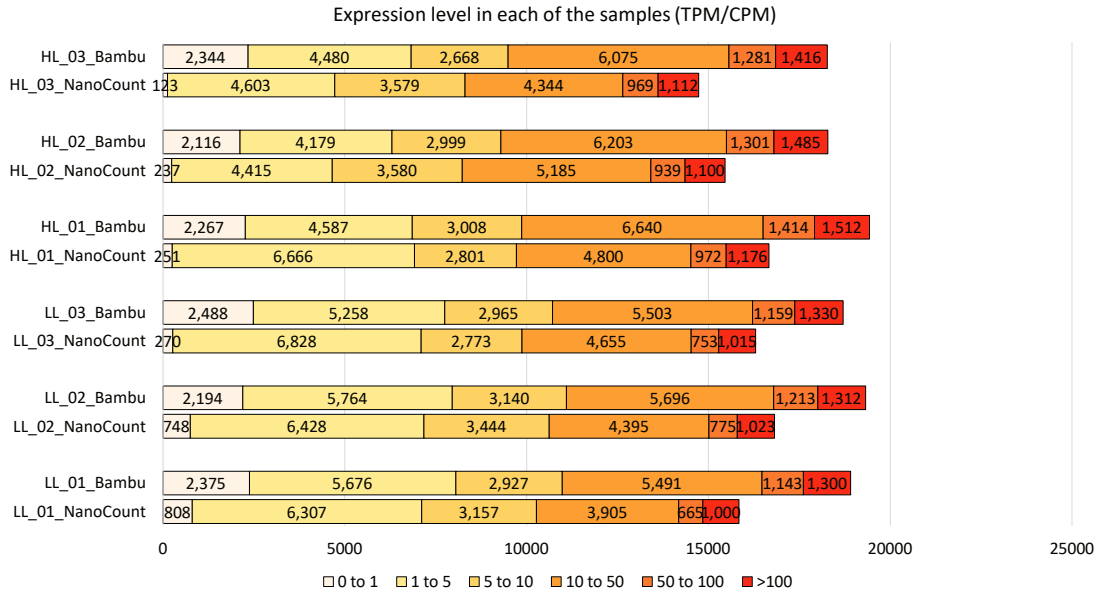


***Figure 4.*** *A comparison of transcript profiling using two different pipelines, NanoCount and Bambu.*

To assess the variability among the samples, principal component analyses (PCA) were performed on the top 5,000 most variable genes identified separately for NanoCount and Bambu pipelines. In general, the sample variance remains relatively consistent between the two pipelines even after the selection of the variability. For 'Bambu', HL samples and LL samples were separated on Dimension 1 (Dim1, **Figure 5A**), which explained 34.5% of the variance. Dimension 2 (Dim2, **Figure 5A**) contributed an additional 19.2% to the variance, capturing a notable deviation of HL_03. By contrast, the Dim1 of the 'NanoCount' dataset accounted for a similar but slightly higher contribution compared to 'Bambu', with 39.6% of the total variance (**Figure 5B**). The Dim2 of the 'NanoCount' dataset explained 19.1% of the variance, which also highlighted the deviation of HL_03 (**Figure 5B**). Therefore, both PCA plots identified distinctions between HL and LL samples, affirming that the measured expression changes align with the underlying biological variations across the samples.

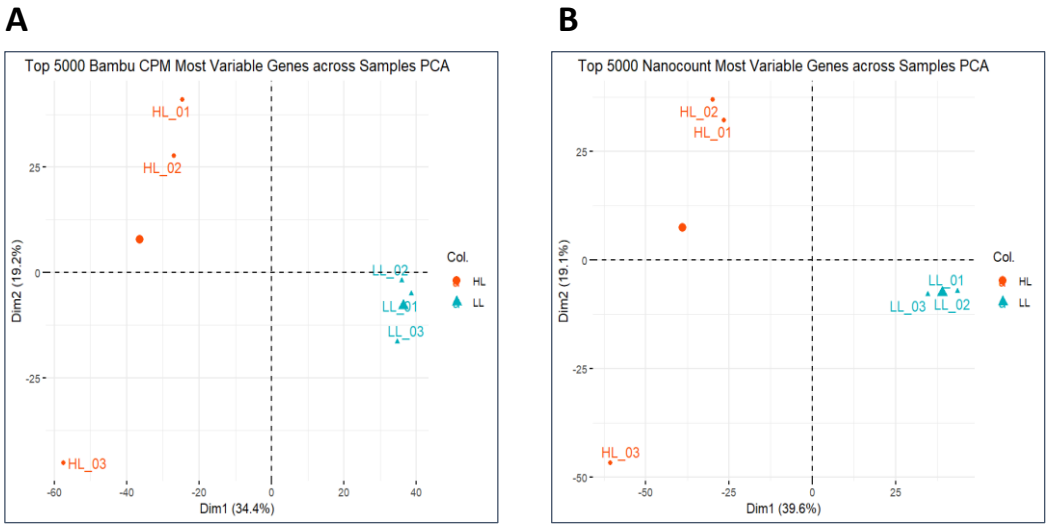


***Figure 5.*** *Principal component analysis (PCA) of HL and LL samples.* ***A,*** *PCA plot of top 5,000 most variable genes quantified by 'Bambu';* ***B,*** *PCA plot of top 5,000 most variable genes quantified by 'NanoCount'.*

## 3.5. Differential Gene Expression Analysis Derived from Full-length cDNA Sequencing Data

By comparing expression data between HL and LL samples, 776 and 846 DEGs were identified from the "Bambu" and "NanoCount" datasets, respectively (|fold change| ≥ 2, adjusted *p*-value < 0.05, **Figure 6A-D**). In total, 302 genes were overlapped in both datasets, 544 DEGs were detected only in the 'NanoCount' dataset, and 474 DEGs were exclusively identified in the 'Bambu' dataset (**Figure 6C**). In the 'Bambu' dataset, 335 genes exhibited up-regulation, and 441 genes displayed down-regulation. Meanwhile, in the 'NanoCount' dataset, there were 369 up-regulated genes and 477 down-regulated genes (**Figure 6D**). For both datasets, the DEGs exhibited distinctively opposite expression patterns comparing HL and LL conditions (**Figure 6E** and **F**).

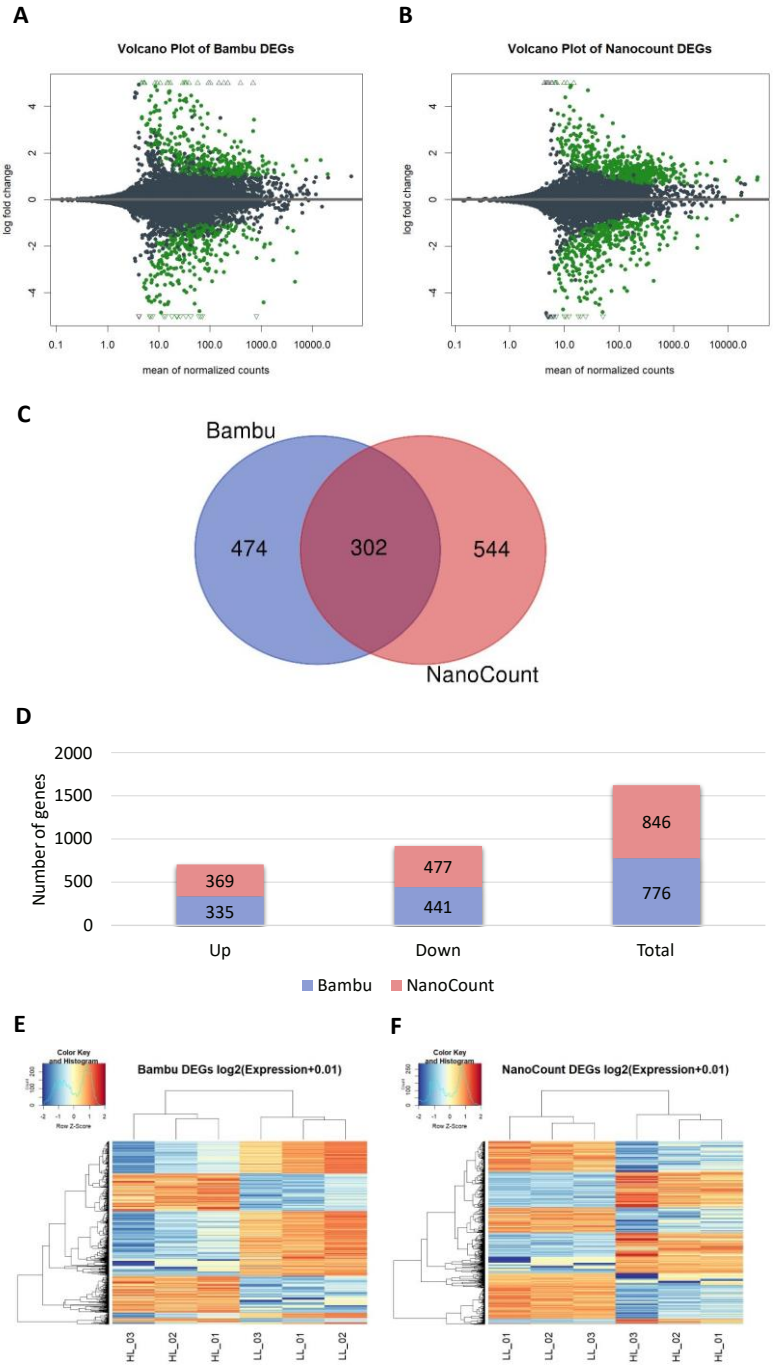

***Figure 6.*** *DEGs analysis.* ***A,*** *the volcano plot illustrating the distribution of DEGs (green color) from 'Bambu' dataset;* ***B,*** *the volcano plot illustrating the distribution of DEGs (green color) from 'NanoCount' dataset;* ***C,*** *the Venn diagram depicting the overlap between 'Bambu' and 'NanoCount' DEGs;* ***D,*** *statistics of DEGs from 'Bambu' and 'NanoCount'. Significant down and up DEGs were determined by |fold change| ≥2 and adjusted p-value < 0.05;* ***E,*** *the heatmap showing the expression pattern of all DEGs detected from 'Bambu' dataset. The heatmap is clustering the log2 (expression data+ 0.01) of each gene;* ***F,*** *heatmap clustering analysis on the expression pattern of the DEGs detected from 'NanoCount' dataset. The heatmap is clustering the log2 (expression data+ 0.01) of each gene.*

## 3.6. Gene Ontology (GO) Term Enrichment of DEGs derived from Full-length cDNA Sequencing Data

To delve into the biological processes triggered by the long-term HL stress, I first employed BLAST analysis to map the *H. incana* DEG sequences to their *A. thaliana* homologs. This allowed me to utilize the abundant functional annotation data available from the model plant species. Subsequently, I analyzed the GO terms and KEGG pathways to assess their functional enrichment. For both pipelines, the top 20 most significant HL-induced GO enrichments were determined by the gene ratio and the adjusted false discovery rate (FDR) (**Figure 7**). For HL down-regulated DEGs, both datasets demonstrated a pronounced enrichment in GO terms related to rRNA biogenesis and assembly. Additionally, GO terms related to translation, peptide biosynthesis, organelle assembly, and amide biosynthetic process were also enriched in both datasets. Further exploration of the 'Bambu' dataset highlighted exclusive enrichment in GO terms associated with organonitrogen compound biosynthesis and cytoplasmic translation. In contrast, 'NanoCount' dataset featured unique enrichment in GO terms related to photosynthesis, chlorophyll biosynthesis, tetrapyrrole biosynthesis, porphyrin-containing compound biosynthesis, and adaxial/abaxial pattern specification. For HL up-regulated DEGs, GO terms related to starch metabolism, response to temperature stimulus, response to oxygen-containing compound, response to the metal ion, response to inorganic substance, glyoxylate and dicarboxylate metabolism, and detoxification were enriched in both datasets. Examining upregulated DEGs in the 'Bambu' dataset unveiled exclusive enrichment in GO terms associated with rhythmic process, response to virus, response to radiation, response to oxidative stress, response to light stimulus, response to cold, response to carbohydrate, response to blue light, precursor metabolites and energy generation, circadian rhythm, and cellular catabolic process. Conversely, 'NanoCount' dataset exhibited distinctive enrichment in GO terms related to regulation of photosynthesis, flavonoid biosynthesis, toxin catabolic process, sulfur compound metabolism, response to water deprivation, response to cadmium ion, glutathione metabolism, and several amino-acids (glycine, serine, threonine, alanine, aspartate, glutamate, and alfa-amino acid) metabolism.

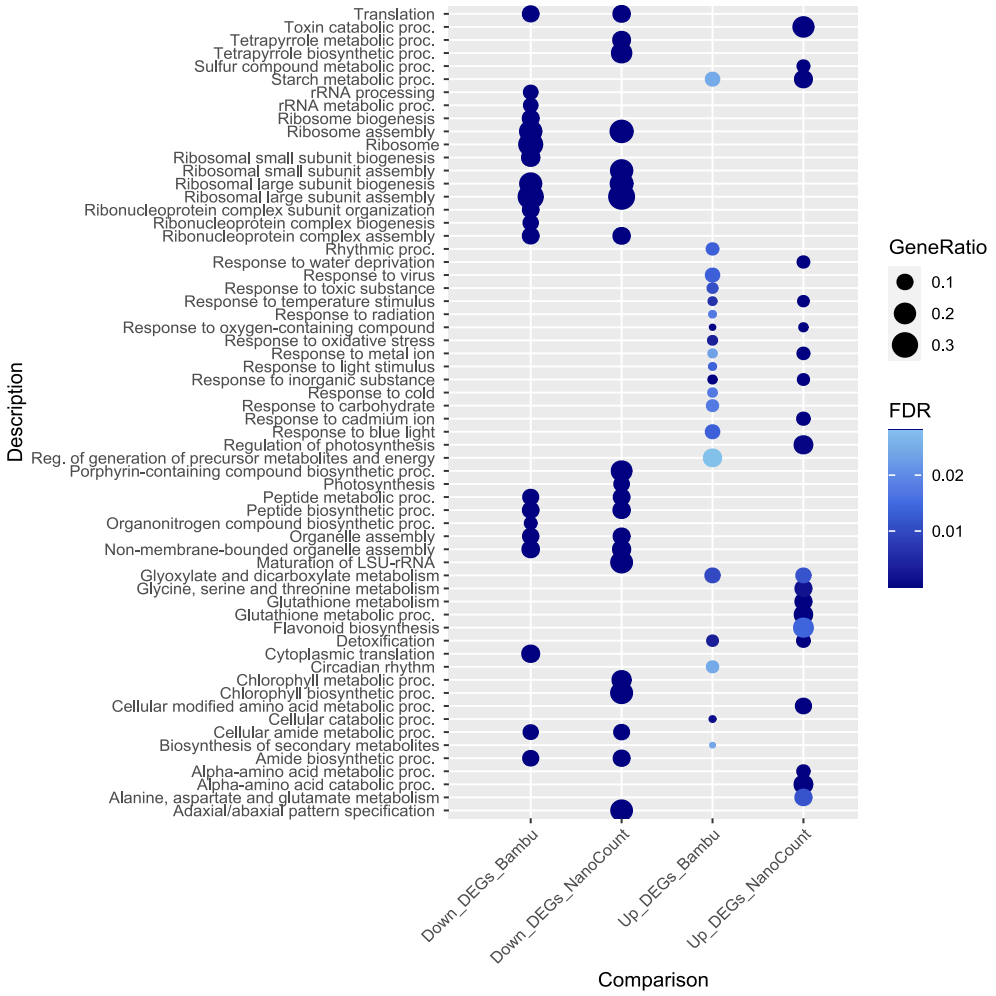


***Figure 7.*** *Comparison of top 20 high-light induced GO enrichment items between 'Bambu' and 'NanoCount' datasets.*

In comparing the sets of DEGs outlined in **Figure 6C** and the top GO terms presented in **Figure 7**, it becomes apparent that the 'NanoCount' dataset offers a more comprehensive and biologically informative perspective than the 'Bambu' dataset. As a result, I have elected to utilize the 'NanoCount' dataset for subsequent downstream analyses.

In total, 338 and 351 *A. thaliana* homologs were identified for the up-regulated and down-regulated *H. incana* genes, respectively. Furthermore, during the co-expression network analysis, several hub-genes centralize other genes with functional and biological associations, formulating distinct subgroups (**Figure 8A** and **B**). For the up-regulated GO terms, three subgroups are illustrated on the general bitmap (**Figure 8A**). Beginning with *P5CSA,* a hub-gene with the most connection nodes. *P5CSA* connects various genes involved in the plant's response to water deprivation, salt stress, and abscisic acid. Moving to *GPX6* (*GLUTATHIONE PEROXIDASE 6*, AT4G11600), another hub-gene involved in cellular oxidant detoxification, integrating genes associated with multiple detoxification processes. Moreover, the hub-gene *HSP70-4* (*HEAT SHOCK PROTEIN 70*, *AT3G12580*), not only centralizes the genes responsive to heat stress but also exhibits close clustering with genes (*CHALCONE--FLAVANONE ISOMERASE 3*, *CALRETICULIN-3*, *FLAVANONE 3-HYDROXYLASE, etc.*) implicated in the response to light stimulus, photoinhibition, and flavonoid biosynthesis. Notably, between the subgroups, I also observed co-expression of specific genes (e.g., *CATALASE-2*, *GLUTATHIONE S-TRANSFERASE F12*, *CALMODULIN-7*) with the hub-genes, thereby facilitating interlinking among these subgroups (**Figure 6A**). In contrast, the down-regulated

GO terms can be primarily categorized into two distinct subgroups (**Figure 8B**). *RPL13* (*50S RIBOSOMAL PROTEIN L13*, AT1G78630) emerged as a centralizing force, orchestrating genes associated with rRNA activities, while *PPL1* (*PSBP-LIKE PROTEIN 1*, AT3G55330) played a pivotal role in grouping genes related to photosynthesis, chlorophyll metabolism, and chloroplast development.

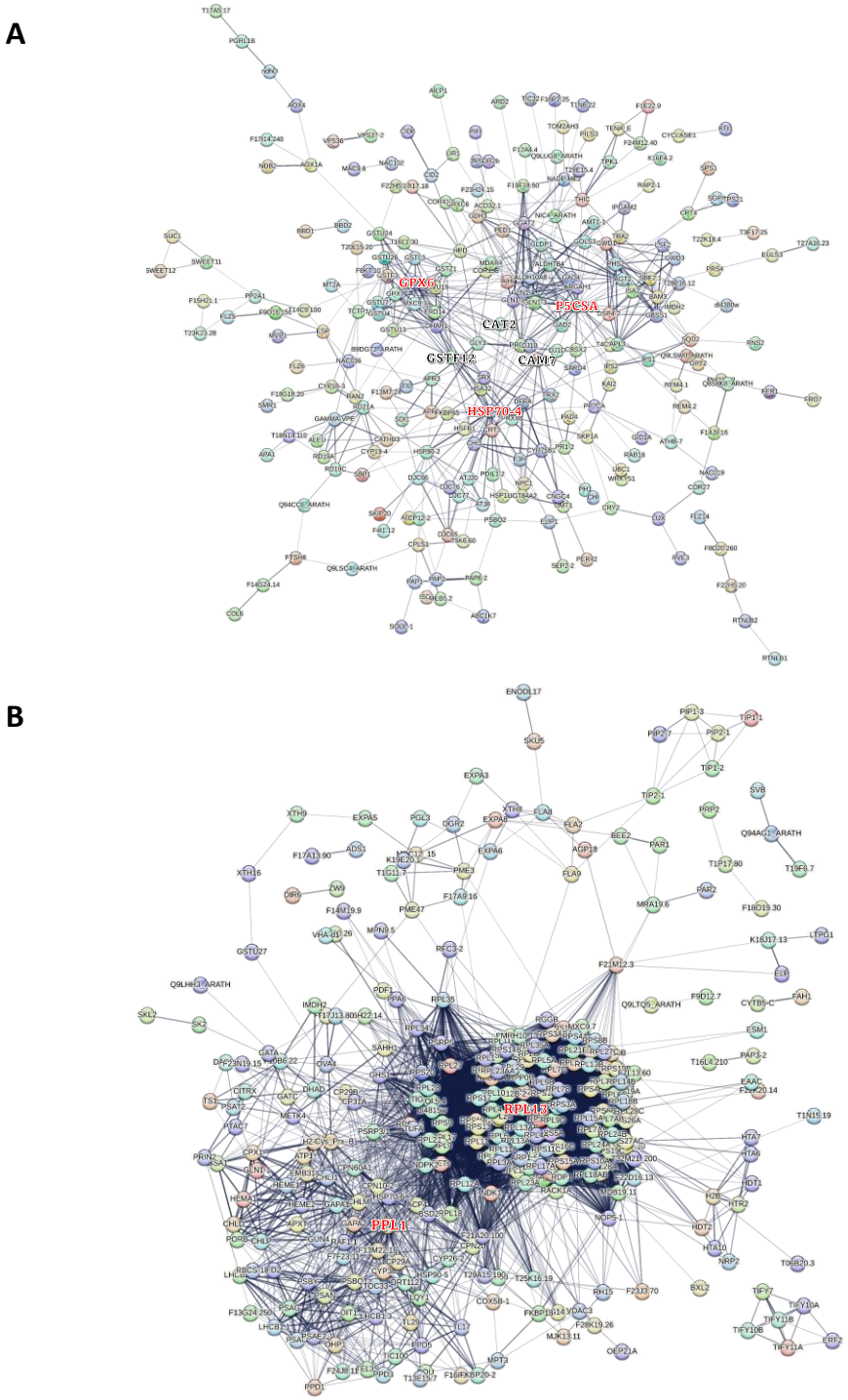


***Figure 8.*** *Comparison of co-expression networks of HL up-regulated and down-regulated genes.* ***A,*** *co-expression network of HL up-regulated genes. P5CSA, hub-gene of the water deprivation-related genes subgroup. GPX6, hub-gene of the detoxification-related genes subgroup. HSP70-4, hub-gene of the heat-stress and light-stress response related genes subgroup.* ***B,*** *a co-expression network of HL down-regulated genes. RPL13, hub-gene of the ribosomal RNA activities related genes subgroup. PPL1, hub-gene of the photosynthesis-related genes subgroup. Networks were generated based on STRING database v12 (Szklarczyk et al., 2020).*

## 3.7. Full-length cDNA Sequencing Facilitated the Detection of Differentially Expressed Gene Copies and Isoforms in the *H. incana* transcriptome

To further investigate the identified DEGs, their log2-transformed expression data (with a pseudo-count of 0.01) for their homologous *H. incana* genes were visualized in heatmaps (**Figure 9A** and **B**). The heatmaps revealed distinct clusters and expression profiles, indicating potential functional divergence among the gene copies. Upon closer inspection, it became evident that several up-regulated DEGs had multiple copies that were either not expressed, conversely expressed, or maintained similar expression levels to the low-light samples, underscoring the existence of differentially expressed copies within the same gene family. To exemplify, I selected a specific example for detailed examination (**Figure 9C**). *HINC_NIJ_V1_G0202360* and *HINC_NIJ_V1_G0091340*, a pair of well-conserved copies, mapped to the same Arabidopsis homolog (*PEROXIREDOXIN-2B*, At1g65980) whereas displayed varied expression responses. *HINC_NIJ_V1_G0202360* was up-regulated, while *HINC_NIJ_V1_G0091340* exhibited reduced expression levels as observed under HL treatment. In terms of isoform expression, 899 differentially expressed isoforms were identified. Remarkably, the expression heatmaps unveiled intricate variations within genes, as individual isoforms demonstrated distinct expression dynamics **(Figure 9D** and **E**). For instance, the gene *HINC_NIJ_V1_G0263230* exhibited three isoforms, while only one of them (*HINC_NIJ_V1_G0263230.1*) emerged as significantly up-regulated under HL stress (**Figure 9F**).

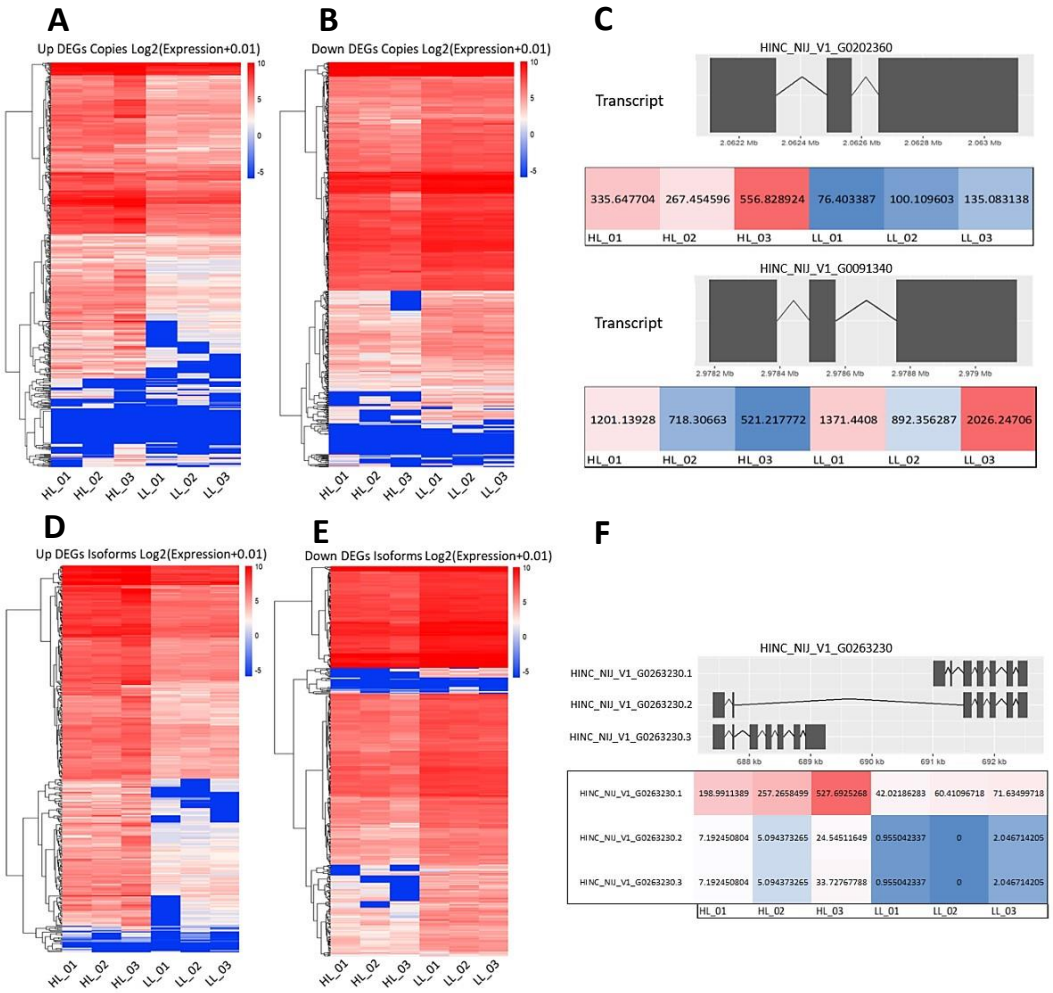

***Figure 9.*** *Differentially expressed gene copies and isoforms.* ***A,*** *the heatmap displaying the expression patterns of all identified up-regulated DEGs' copies in H. incana;* ***B,*** *the heatmap displaying the expression patterns of all identified down-regulated DEGs' copies in H. incana;* ***C,*** *the example of differentially expressed H.incana copies, HINC_NIJ_V1_G0202360 and HINC_NIJ_V1_G0091340;* ***D,*** *the heatmap displaying the expression patterns of all identified up-regulated DEGs' isoforms in H. incana;* ***E,*** *the heatmap displaying the expression patterns of all identified down-regulated DEGs' isoforms in H. incana;* ***F,*** *the example of differentially expressed isoforms of one H.incana gene, HINC_NIJ_V1_G0263230. The heatmaps are clustering the log2(expression data + 0.01) of each copy/isoform. 'HL_01-03', three HL biological replications. 'LL_01-03', three LL biological replications.*

Finally, I have identified 10 genes among the differentially expressed isoforms that exhibit strong associations with HL-induced responses, thus warranting further investigation for gene functional analysis (**Figure 10**). These candidates encompass genes involved in photosynthesis (*PPL1*, *PSBP DOMAIN-CONTAINING PROTEIN 1, PGR5-LIKE PROTEIN 1B* and *RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1B*), response to light stimulus (*TRAF-LIKE FAMILY PROTEIN F9D12.7*, *PROTEIN PLASTID TRANSCRIPTIONALLY ACTIVE 7*), response to oxidative stress (*GLUTAREDOXIN-C5* and *THYLAKOID LUMENAL 29 KDA PROTEIN*), cell development (*CYCLASE-LIKE PROTEIN 2*), and flavonoid biosynthesis (*F3H*).

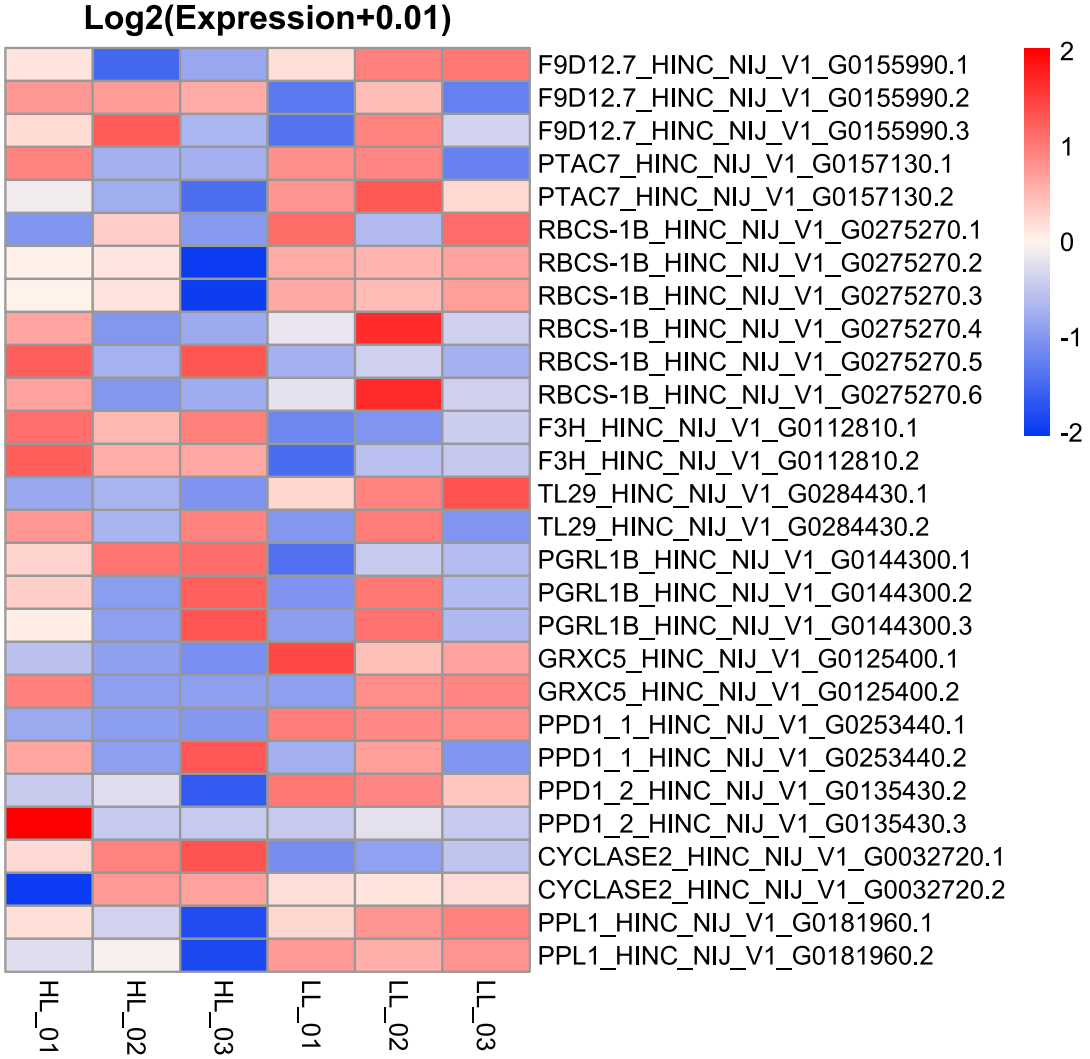


***Figure 10.*** *The heatmap illustrates the expression of 10 candidate isoforms across HL and LL samples. The expression levels are represented using the log2 transformation (expression data + 0.01) for each isoform.*

# 4. Discussion

## 4.1. Long-Term Light Stress Adaptive Response in *A. thaliana*

The long-term high-light experiment provides insights into sustained molecular responses to HL stress, revealing a significant impact on the plant transcriptome. Integration of long-term HL gene expression profiles with the short-term experiment uncovers intriguing patterns, emphasizing shared and distinct regulatory elements among 1,593 overlapped DEGs. Notably, consistent regulation in 711 DEGs, reversal in 72 DEGs, and restoration in 810 genes suggest a dynamic interplay of responses.

The observed decrease in the expression of photosynthetic genes under both short-term and long-term HL conditions signifies a conserved responsive biology. Similarly, the persistence of down-regulated responses to radiation, light intensity, and light stimulus underscores the temporal stability of certain stress-related pathways. However, the distinct patterns in long-term exposure, with several photosynthesis-related genes showing restoration from down-regulation, indicate a nuanced adaptation to prolonged stress. Notably, the intriguing observation of altered photosystem I (PSI) gene expression, transitioning from down-regulation to a normal state in response to long-term HL stress, aligns with the photoprotective role of PSI. As suggested by a previous study on *Panax notoginseng*, cyclic electron flow around PSI has been shown to safeguard both PSI and PSII under long-term HL stress (Cun et al., 2023). In this context, the altered expression of PSI genes could be viewed as a molecular adaptation aimed at sustaining the photosynthesis ability under prolonged HL exposure.

Robust reactions to ROS were evident in both short-term and long-term HL conditions, emphasizing the crucial role of oxidative stress management. However, the restoration of some up-regulated oxidative stress-responsive genes under prolonged exposure to HL suggests an adaptive fine-tuning over time. Remarkably, this process coincided with the restoration from the up-regulation of diverse genes involved in flavonoid biosynthesis. It is widely recognized that anthocyanin, as a member of flavonoids, belongs to ROS-inducible compounds and plays a crucial role in modulating the sensitivity toward ROS stresses (Khoo et al., 2017; Xu et al., 2017). Consequently, I postulate that *A. thaliana* plants may exhibit altered strategies to cope with ROS under prolonged light stress, which could manifest as either reduced sensitivity or the activation of alternative responsive pathways. This speculation is supported to some extent by the fact that the up-regulation of genes associated with isoprenoid activity under long-term HL condition compared to their down-regulation under short-term HL condition.

Distinct mechanisms in response to drought were also highlighted, with long-term exposure stimulating drought tolerance responses, while short-term exposure induces more immediate responses to water deprivation and salt stress. A few studies have described the mechanisms associated with the two stressors, such as the light-dependent stomata closure affecting the plant drought tolerance (Mao et al., 2005). The short-term experiment mentioned the identification of DEGs related to drought but did not discuss it in detail. Nonetheless, I am conservative about the water-deficit responses observed in *A. thaliana* plants during the short-term experiment, as they employed continuous cold wind beneath the plants to separate heat stress from HL conditions. It is important to note that windy air contributes to water evaporation. On the other hand, the long-term HL data exhibited significant up-regulation in the alpha-linolenic acid metabolism pathway that aligns with findings from Zi et al. (2022), who observed similar up-regulation patterns in response to long-term drought stress in maize (*Zea mays* L.) (Zi et al., 2022). This consistency across different plant species underscores the conservation of certain stress-responsive pathways and indicates the cross-talk between light-stress and drought-stress. , efficient transport of sulfate has been demonstrated to play a crucial role in plant

responses to drought and salinity (Gallardo et al., 2014), thereby highlighting its significance in adapting to long-term stress conditions.

Moreover, the restoration of up-regulated ABA- and JA-signaling observed in the long-term HL experiment aligns with several previous studies, in which ABA and JA have been suggested to be involved in short-term light acclimation (Dietz, 2015; Huang et al., 2019; Ramel et al., 2013; Suzuki et al., 2013). Therefore, it is more plausible that ABA and JA biosynthesis are functioning in stress response specifically after exposure to HL for a short time.

## 4.2. Illumina Short-read Dataset Provides Insight into Long-Term Light Stress Adaptive Response in *H. incana*

The Illumina dataset provides valuable insight into the biological response of *H. incana* to long-term HL condition. On the one hand, the down-regulation of genes associated with ribosome and rRNA activities is particularly noteworthy. Sulpice et al. (2014) discovered that Proteaceae species exhibit a reduction in rRNA abundance as an adaptive strategy to enhance phosphorus use efficiency, thereby enabling them to maintain a relatively high photosynthesis rate even under conditions of limited phosphorus availability (Sulpice et al., 2014). Therefore, I speculated that *H. incana* utilizes rRNA suppression to maintain basic cellular functions and protein synthesis, manifesting a prioritization of resources away from general growth and development under HL stress. Similarly, the down-regulation of translation and DNA replication implies a concerted effort to conserve energy and resources under prolonged HL stress. In addition, the down-regulation of porphyrin and chlorophyll metabolism indicates negative regulation of photosynthesis as they are involved in the critical energy conversion step of photosynthesis (Mauzerall, 1977). On the other hand, the observed up-regulation of amino acid metabolic pathways suggests a potential role in providing energy or building blocks for stress-related processes. Arginine, proline, valine, leucine, and isoleucine are widely recognized growth-necessary amino acids that also play an important role in responding to various stressors such as drought, salinity, and osmotic stress (Batista-Silva et al., 2019; Hussein et al., 2022; Ingrisano et al., 2023). Therefore, it is plausible that these amino acids serve as precursors for the synthesis of protective molecules or signaling compounds involved in the adaptive response to HL stress in *H. incana*. Additionally, in conjunction with the observed up-regulation of stress response pathways, encompassing drought, salt stress, oxidative stress, and hypoxia, my analysis underscores the adaptability of *H. incana* to long-term light-stress conditions, as previously identified in *A. thaliana*. However, the differentially expressed copies of the top DEGs within these stress-responsive pathways distinguish *H. incana* from its relatives. These variations in gene expression profiles, although not directly indicative of their functional contribution, point towards an association between CNVs and HL adaptation in H. incana that aligns with the findings reported by Garassino et al. (2022). Unraveling the functional significance of these differentially expressed copies represents an exciting avenue for future research, offering the potential to uncover specific molecular adaptations that contribute to *H. incana*'s resilience in the face of long-term HL stress.

## 4.3. Nanopore Long-read cDNA Sequencing Reveals Isoform-level Adaptive Responses to Long-term HL Stress in *H. incana*.

Nanopore long-read sequencing emerged as a powerful tool for acquiring full-length transcripts, yielding a comprehensive dataset of approximately 12Gb of cDNA data from six samples.

A comprehensive evaluation of key sequencing parameters, including read quality, count, and length, was conducted. Despite variability in read counts, ranging from over 4 million to approximately 2.7 million, a consistent pattern of read quality and length emerged. Notably, the N50 around 800 bp underscores the effectiveness of Nanopore long-read sequencing in capturing substantial portions of full-length transcripts; however, the adequacy of the sequencing depth for thorough analysis remains uncertain at this juncture. Further assessment and consideration of depth-related factors are warranted to ensure the reliability and completeness of subsequent analyses.

Regarding gene expression quantification, the two pipelines (‘NanoCount’ and ‘Bambu’) are relatively comparable since full-length transcripts were generated, and both normalization strategies (TPM and CPM) ensured comparability across samples, accounting for variations in library size and enabling meaningful cross-sample comparisons. Subsequent PCA analysis was performed on the top 5,000 most variable genes, where both pipelines consistently identified distinctions between HL and LL samples, affirming that the measured expression changes accurately reflect underlying biological variations across the samples. The robustness and reproducibility of these results across different pipelines enhance the confidence in the reliability of my gene expression quantification. The identification of 776 DEGs by "Bambu" and 846 DEGs by "NanoCount" underscores the sensitivity of these pipelines to capture the subtleties of gene expression changes. The overlap of 302 genes suggests a core set of DEGs, implying shared responses to the experimental conditions. However, the identification of unique genes by each pipeline ("NanoCount" with 544 and "Bambu" with 474) emphasizes the need for careful examination to understand the specific characteristics revealed by each methodology. To be more specific, the 'Bambu' dataset highlighted pathways associated with organonitrogen compound biosynthesis and cytoplasmic translation, suggesting specific metabolic adaptations. In contrast, the 'NanoCount' showcased unique enrichments related to photosynthesis, chlorophyll biosynthesis, tetrapyrrole biosynthesis, and adaxial/abaxial pattern specification, shedding light on additional facets of *H. incana*'s response that might be overlooked in a singular approach. The choice to utilize the 'NanoCount' dataset for further analyses stems from its capacity to encompass a wider range of responses, affording a more comprehensive understanding of intricate adaptations to long-term HL exposure in *H. incana*.

Employing *A. thaliana* homologs, several key gene groups were identified, highlighting the central roles of hub-genes in deciphering *H. incana*'s adaptive strategies under light stress. Among the up-regulated DEGs, *P5CSA* emerged as a central hub-gene, crucial for water deprivation, salt stress, and ABA responses, echoing its significance in the Illumina dataset and in *A. thaliana* studies (Funck et al., 2020; Mahmud et al., 2023). Another hub-gene, *GPX6*, vital for oxidant detoxification, demonstrated concerted efforts against oxidative stress, as indicated in previous research on the *A. thaliana GPX* gene family (Brigelius-Flohé & Maiorino, 2013; Passaia et al., 2014; Xiang & Hou, 2021). Similarly, the hub-gene *HSP70-4* centralized genes that are responsive to heat stress while being closely clustered with those involved in light-related processes, highlighting interconnected responses to HL exposure. Co-expression of specific genes (e.g., *CAT2*, *GSTF12*, *CAM7*) with hub-genes underscored collaborative efforts, indicating a synergistic response to light-stress conditions. Although Huang et al. (2019) argued that the accuracy of the comprehension regarding transcriptional regulation associated with

light-driven signals is compromised by such light-dependent heat stress, it's undeniable that light is a source of heat (Huang et al., 2019; Katzin et al., 2023). Therefore, this study contends that investigating the HL stress adaptation mechanism of *H. incana* under natural conditions would be better served by maintaining a harmonious interplay between light and heat. Concerning down-regulated DEGs, the down-regulation of photosynthesis aligns with the short-term HL experiment on A. thaliana (Huang et al., 2019) and the Illumina data of *H. incana*. It indicates a conserved responsive mechanism of modulating the photosynthetic apparatus to prevent photodamage and enhance the resilience to light-stress condition (Kasahara et al., 2002; Ruban, 2016).

More importantly, by utilizing Nanopore long-read sequencing, the investigation is moving beyond gene copies, the exploration of differentially expressed isoforms further enriches our understanding of gene expression dynamics. The identification of 899 differentially expressed isoforms unveils a level of complexity that goes beyond the gene level. I've selected 10 candidate genes among the differentially expressed isoforms, associated with HL-induced responses. The diverse functions of these candidate genes, ranging from photosynthesis and response to light stimulus to oxidative stress and flavonoid biosynthesis, suggest a multifaceted response to HL stress. These candidates provide a valuable starting point for further investigation, encouraging further exploration to elucidate their specific roles in the adaptive mechanisms of *H. incana* under HL conditions.

# 5. Conclusion

In conclusion, the findings of this study strongly support the initial hypothesis, revealing compelling evidence of differential expression among gene copies and/or isoforms related to photosynthesis under long-term HL stress. The study on *A. thaliana*'s response to long-term HL reveals a complex and dynamic adaptation at the molecular level. The significant number of differentially expressed genes, particularly in photosynthetic pathways and stress-related responses, underscores the substantial impact of prolonged HL exposure on the plant's transcriptome. Noteworthy findings include the nuanced regulation of photosynthesis-related genes, the adaptive fine-tuning of oxidative stress-responsive genes over time, and the restoration of hormonal signaling pathways. The observed enhancement of drought tolerance mechanisms, such as the regulation of sulfate transmembrane transport, further highlights the plant's ability to adapt to extended stress conditions. The investigation into *H. incana*'s response to long-term HL provides insights at both the gene and isoform levels. The down-regulation of genes associated with rRNA activities suggests a strategic suppression to enhance phosphorus use efficiency, prioritizing resources away from general growth and development under HL stress. Concurrently, the up-regulation of amino acid metabolic pathways, particularly for growth-necessary amino acids, implies their potential role in providing energy or building blocks for stress-related processes. The observed down-regulation of porphyrin and chlorophyll metabolism indicates a negative regulation of photosynthesis, while the up-regulation of stress response pathways underscores the specific adaptability of *H. incana* to prolonged light stress. Nanopore long-read cDNA sequencing reveals isoform-level adaptive responses. The identification of differentially expressed isoforms associated with various stress-related processes highlights the multifaceted nature of adaptation to HL stress in *H. incana*. Hub-gene analysis reveals pivotal genes, such as *P5CSA*, *GPX6*, and *HSP70-4*, which play central roles in integrating stress-related pathways and detoxification processes. Overall, these findings contribute to a comprehensive understanding of plant resilience in the face of long-term HL stress, paving the way for further research into specific molecular adaptations and isoform-specific roles in HL adaptation.

# 6. Recommendations

Based on the findings presented in this study, there are three key recommendations for further research:

## 6.1. Gene Function Validation through CRISPR-Cas Knock-out/Knock-down

The study has identified several candidate isoforms associated with adaptive responses to long-term HL stress in *H. incana*. To validate the functional significance of these genes, it is highly recommended to employ CRISPR-Cas gene editing techniques for knock-out or knock-down experiments. By selectively modifying the expression of key genes, we can directly assess the specific roles played by isoforms in the HL adaptation of *H. incana*.

## 6.2. Direct RNA Sequencing

Direct RNA sequencing methods should be considered to avoid the need for reverse transcription in generating cDNA. The utilization of the direct RNA sequencing technique confers the advantage of capturing complete transcripts without the biases introduced during reverse transcription and PCR. This approach provides a more accurate representation of transcript isoforms and improves the reliability of transcriptome characterization.

## 6.3. Increase Sequencing Depth for Enhanced Data Capture

To further enrich the dataset and capture a more comprehensive view of the transcriptome, it is recommended to increase the sequencing depth. A higher sequencing depth will result in more reads, providing greater coverage of the transcriptome and enhancing the sensitivity for detecting rare or low-abundance transcripts and isoforms. These rare or low-abundance transcripts may play crucial roles in the adaptive responses of *H. incana* to HL stress, thus contributing to a more detailed understanding of gene expression dynamics and improving the reliability of differential expression analyses.

# Reference


Apel, K., & Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.*, *55*, 373-399.

Aro, E.-M., Virgin, I., & Andersson, B. (1993). Photoinhibition of photosystem II. Inactivation, protein damage, and turnover. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1143*(2), 113-134.

Batista-Silva, W., Heinemann, B., Rugen, N., Nunes-Nesi, A., Araújo, W. L., Braun, H. P., & Hildebrandt, T. M. (2019). The role of amino acid metabolism during abiotic stress release. *Plant Cell Environ*, *42*(5), 1630-1644. https://doi.org/10.1111/pce.13518

Bela, K., Horváth, E., Gallé, Á., Szabados, L., Tari, I., & Csiszár, J. (2015). Plant glutathione peroxidases: Emerging role of the antioxidant enzymes in plant development and stress responses. *Journal of Plant Physiology*, *176*, 192-201. https://doi.org/https://doi.org/10.1016/j.jplph.2014.12.014

Bernard, E., Jacob, L., Mairal, J., & Vert, J.-P. (2014). Efficient RNA isoform identification and quantification from RNA-Seq data with network flows. *Bioinformatics*, *30*(17), 2447-2455.

Brigelius-Flohé, R., & Maiorino, M. (2013). Glutathione peroxidases. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *1830*(5), 3289-3303.

Canvin, D. T., Berry, J. A., Badger, M. R., Fock, H., & Osmond, C. B. (1980). Oxygen exchange in leaves in the light. *Plant Physiology*, *66*(2), 302-307.

Chen, Y., Sim, A., Wan, Y. K., Yeo, K., Lee, J. J. X., Ling, M. H., Love, M. I., & Göke, J. (2023). Context-aware transcript quantification from long-read RNA-seq data with Bambu. *Nature methods*, 1-9.

Cock, P. J. A., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015). NCBI BLAST+ integrated into Galaxy. *GigaScience*, *4*(1). https://doi.org/10.1186/s13742-015-0080-7

Cun, Z., Xu, X.-Z., Zhang, J.-Y., Shuang, S.-P., Wu, H.-M., An, T.-X., & Chen, J.-W. (2023). Responses of photosystem to long-term light stress in a typically shade-tolerant species Panax notoginseng. *Frontiers in Plant Science*, *13*, 1095726.

Devireddy, A. R., Zandalinas, S. I., Gómez-Cadenas, A., Blumwald, E., & Mittler, R. (2018). Coordinating the overall stomatal response of plants: Rapid leaf-to-leaf communication during light stress. *Science Signaling*, *11*(518), eaam9514. https://doi.org/doi:10.1126/scisignal.aam9514

Dietz, K.-J. (2015). Efficient high light acclimation involves rapid processes at multiple mechanistic levels. *Journal of Experimental Botany*, *66*(9), 2401-2414. https://doi.org/10.1093/jxb/eru505

Feng, L., Raza, M. A., Li, Z., Chen, Y., Khalid, M. H. B., Du, J., Liu, W., Wu, X., Song, C., & Yu, L. (2019). The influence of light intensity and leaf movement on photosynthesis characteristics and carbon balance of soybean. *Frontiers in Plant Science*, *9*, 1952.

Filichkin, S. A., Hamilton, M., Dharmawardhana, P. D., Singh, S. K., Sullivan, C., Ben-Hur, A., Reddy, A. S., & Jaiswal, P. (2018). Abiotic stresses modulate landscape of poplar transcriptome via alternative splicing, differential intron retention, and isoform ratio switching. *Frontiers in Plant Science*, *9*, 5.

Fini, A., Brunetti, C., Di Ferdinando, M., Ferrini, F., & Tattini, M. (2011). Stress-induced flavonoid biosynthesis and the antioxidant machinery of plants. *Plant Signaling & Behavior*, *6*(5), 709-711. https://doi.org/10.4161/psb.6.5.15069

Funck, D., Baumgarten, L., Stift, M., Von Wirén, N., & Schönemann, L. (2020). Differential contribution of P5CS isoforms to stress tolerance in Arabidopsis. *Frontiers in Plant Science*, *11*, 565134.

Gallardo, K., Courty, P. E., Le Signor, C., Wipf, D., & Vernoud, V. (2014). Sulfate transporters in the plant's response to drought and salinity: regulation and possible functions. *Front Plant Sci*, *5*, 580. https://doi.org/10.3389/fpls.2014.00580

Garassino, F., Wijfjes, R. Y., Boesten, R., Reyes Marquez, F., Becker, F. F., Clapero, V., van den Hatert, I., Holmer, R., Schranz, M. E., & Harbinson, J. (2022). The genome sequence of Hirschfeldia incana, a new Brassicaceae model to improve photosynthetic light-use efficiency. *The Plant Journal*, *112*(5), 1298-1315.

Gitelson, A. A., Peng, Y., Arkebauer, T. J., & Suyker, A. E. (2015). Productivity, absorbed photosynthetically active radiation, and light use efficiency in crops: Implications for remote sensing of crop primary production. *Journal of Plant Physiology*, *177*, 100-109.

Gleeson, J., Leger, A., Prawer, Y. D. J., Lane, T. A., Harrison, P. J., Haerty, W., & Clark, M. B. (2021). Accurate expression quantification from nanopore direct RNA sequencing with NanoCount. *Nucleic acids research*, *50*(4), e19-e19. https://doi.org/10.1093/nar/gkab1129

Hastings, P. J., Lupski, J. R., Rosenberg, S. M., & Ira, G. (2009). Mechanisms of change in gene copy number. *Nature Reviews Genetics*, *10*(8), 551-564. https://doi.org/10.1038/nrg2593

Havaux, M. (2014). Carotenoid oxidation products as stress signals in plants. *The Plant Journal*, *79*(4), 597-606. https://doi.org/https://doi.org/10.1111/tpj.12386

He, Z., Ji, R., Havlickova, L., Wang, L., Li, Y., Lee, H. T., Song, J., Koh, C., Yang, J., Zhang, M., Parkin, I. A. P., Wang, X., Edwards, D., King, G. J., Zou, J., Liu, K., Snowdon, R. J., Banga, S. S., Machackova, I., & Bancroft, I. (2021). Genome structural evolution in Brassica crops. *Nature Plants*, *7*(6), 757-765. https://doi.org/10.1038/s41477-021-00928-8

Horton, P., Ruban, A. V., & Walters, R. G. (1994). Regulation of Light Harvesting in Green Plants (Indication by Nonphotochemical Quenching of Chlorophyll Fluorescence). *Plant Physiology*, *106*(2), 415.

Huang, J., Zhao, X., & Chory, J. (2019). The Arabidopsis transcriptome responds specifically and dynamically to high light stress. *Cell Reports*, *29*(12), 4186-4199. e4183.

Hussein, H.-A. A., Alshammari, S. O., Kenawy, S. K. M., Elkady, F. M., & Badawy, A. A. (2022). Grain-Priming with L-Arginine Improves the Growth Performance of Wheat (Triticum aestivum L.) Plants under Drought Stress. *Plants*, *11*(9), 1219. https://www.mdpi.com/2223-7747/11/9/1219

Ingrisano, R., Tosato, E., Trost, P., Gurrieri, L., & Sparla, F. (2023). Proline, Cysteine and Branched-Chain Amino Acids in Abiotic Stress Response of Land Plants and Microalgae. *Plants*, *12*(19), 3410. https://www.mdpi.com/2223-7747/12/19/3410

Kasahara, M., Kagawa, T., Oikawa, K., Suetsugu, N., Miyao, M., & Wada, M. (2002). Chloroplast avoidance movement reduces photodamage in plants. *Nature*, *420*(6917), 829-832. https://doi.org/10.1038/nature01213

Khoo, H. E., Azlan, A., Tang, S. T., & Lim, S. M. (2017). Anthocyanidins and anthocyanins: colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food Nutr Res*, *61*(1), 1361779. https://doi.org/10.1080/16546628.2017.1361779

Kobayashi, N., & DellaPenna, D. (2008). Tocopherol metabolism, oxidation and recycling under high light stress in Arabidopsis. *Plant J*, *55*(4), 607-618. https://doi.org/10.1111/j.1365-313X.2008.03539.x

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094-3100.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lysak, M. A., Koch, M. A., Pecinka, A., & Schubert, I. (2005). Chromosome triplication found across the tribe Brassiceae. *Genome research*, *15*(4), 516-525.

Maai, E., Nishimura, K., Takisawa, R., & Nakazaki, T. (2020). Light stress-induced chloroplast movement and midday depression of photosynthesis in sorghum leaves. *Plant Production Science*, *23*(2), 172-181.

Mahmud, S., Kamruzzaman, M., Bhattacharyya, S., Alharbi, K., Abd El Moneim, D., & Mostofa, M. G. (2023). Acetic acid positively modulates proline metabolism for mitigating PEG-mediated drought stress in Maize and Arabidopsis. *Frontiers in Plant Science*, *14*.

Mauzerall, D. (1977). Porphyrins, Chlorophyll, and Photosynthesis. In A. Trebst & M. Avron (Eds.), *Photosynthesis I: Photosynthetic Electron Transport and Photophosphorylation* (pp. 117-124). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-66505-9_5

Mehrotra, S., & Goyal, V. (2014). Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. *Genomics, proteomics & bioinformatics*, *12*(4), 164-171.

Moustakas, M. (2022). Plant Photochemistry, Reactive Oxygen Species, and Photoprotection. *Photochem*, *2*(1), 5-8. https://www.mdpi.com/2673-7256/2/1/2

Murata, N., Takahashi, S., Nishiyama, Y., & Allakhverdiev, S. I. (2007). Photoinhibition of photosystem II under environmental stress. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1767*(6), 414-421. https://doi.org/https://doi.org/10.1016/j.bbabio.2006.11.019

Passaia, G., Queval, G., Bai, J., Margis-Pinheiro, M., & Foyer, C. H. (2014). The effects of redox controls mediated by glutathione peroxidases on root architecture in Arabidopsis thaliana. *Journal of Experimental Botany*, *65*(5), 1403-1413.

Perumal, S., Koh, C. S., Jin, L., Buchwaldt, M., Higgins, E. E., Zheng, C., Sankoff, D., Robinson, S. J., Kagale, S., Navabi, Z.-K., Tang, L., Horner, K. N., He, Z., Bancroft, I., Chalhoub, B., Sharpe, A. G., & Parkin, I. A. P. (2020). A high-contiguity Brassica nigra genome localizes active centromeres and defines the ancestral Brassica genome. *Nature Plants*, *6*(8), 929-941. https://doi.org/10.1038/s41477-020-0735-y

Powles, S. B. (1984). Photoinhibition of photosynthesis induced by visible light. *Annual review of plant physiology*, *35*(1), 15-44.

Ramel, F., Ksas, B., Akkari, E., Mialoundama, A. S., Monnet, F., Krieger-Liszkay, A., Ravanat, J.-L., Mueller, M. J., Bouvier, F., & Havaux, M. (2013). Light-Induced Acclimation of the Arabidopsis chlorina1 Mutant to Singlet Oxygen *The Plant Cell*, *25*(4), 1445-1462. https://doi.org/10.1105/tpc.113.109827

Ruban, A. V. (2016). Nonphotochemical Chlorophyll Fluorescence Quenching: Mechanism and Effectiveness in Protecting Plants from Photodamage. *Plant Physiology*, *170*(4), 1903-1916. https://doi.org/10.1104/pp.15.01935

Schranz, M. E., Lysak, M. A., & Mitchell-Olds, T. (2006). The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in plant science*, *11*(11), 535-542. https://doi.org/https://doi.org/10.1016/j.tplants.2006.09.002

Siemens, J. (2010). Hirschfeldia. In *Wild Crop Relatives: Genomic and Breeding Resources: Oilseeds* (pp. 171-176). Springer.

Sulpice, R., Ishihara, H., Schlereth, A., Cawthray, G. R., Encke, B., Giavalisco, P., Ivakov, A., Arrivault, S., Jost, R., Krohn, N., Kuo, J., Laliberté, E., Pearse, S. J., Raven, J. A., Scheible, W. R., Teste, F., Veneklaas, E. J., Stitt, M., & Lambers, H. (2014). Low levels of ribosomal RNA partly account for the very high photosynthetic phosphorus-use efficiency of Proteaceae species. *Plant Cell Environ*, *37*(6), 1276-1298. https://doi.org/10.1111/pce.12240

Suzuki, N., Miller, G., Salazar, C., Mondal, H. A., Shulaev, E., Cortes, D. F., Shuman, J. L., Luo, X., Shah, J., Schlauch, K., Shulaev, V., & Mittler, R. (2013). Temporal-Spatial Interaction between Reactive Oxygen Species and Abscisic Acid Regulates Rapid Systemic Acclimation in Plants *The Plant Cell*, *25*(9), 3553-3569. https://doi.org/10.1105/tpc.113.114595

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2020). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, *49*(D1), D605-D612. https://doi.org/10.1093/nar/gkaa1074

Takahashi, S., & Badger, M. R. (2011). Photoprotection in plants: a new light on photosystem II damage. *Trends in plant science*, *16*(1), 53-60. https://doi.org/https://doi.org/10.1016/j.tplants.2010.10.001

Taylor, G., Garassino, F., Aarts, M. G., & Harbinson, J. (2023). Improving C3 photosynthesis by exploiting natural genetic variation: Hirschfeldia incana as a model species. *Food and Energy Security*, *12*(1), e420.

Terashima, I., Hanba, Y. T., Tazoe, Y., Vyas, P., & Yano, S. (2005). Irradiance and phenotype: comparative eco-development of sun and shade leaves in relation to photosynthetic CO2 diffusion. *Journal of Experimental Botany*, *57*(2), 343-354. https://doi.org/10.1093/jxb/erj014

Town, C. D., Cheung, F., Maiti, R., Crabtree, J., Haas, B. J., Wortman, J. R., Hine, E. E., Althoff, R., Arbogast, T. S., Tallon, L. J., Vigouroux, M., Trick, M., & Bancroft, I. (2006). Comparative Genomics of Brassica oleracea and Arabidopsis thaliana Reveal Gene Loss, Fragmentation, and Dispersal after Polyploidy. *The Plant Cell*, *18*(6), 1348-1359. https://doi.org/10.1105/tpc.106.041665

Turner, D. P., Urbanski, S., Bremer, D., Wofsy, S. C., Meyers, T., Gower, S. T., & Gregory, M. (2003). A cross-biome comparison of daily light use efficiency for gross primary production. *Global Change Biology*, *9*(3), 383-395. https://doi.org/https://doi.org/10.1046/j.1365-2486.2003.00573.x

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.-H., Bancroft, I., Cheng, F., Huang, S., Li, X., Hua, W., Wang, J., Wang, X., Freeling, M., Pires, J. C., Paterson, A. H., Chalhoub, B., . . . The Brassica rapa Genome Sequencing Project, C. (2011). The genome of the mesopolyploid crop species Brassica rapa. *Nature genetics*, *43*(10), 1035-1039. https://doi.org/10.1038/ng.919

Xiang, D., & Hou, X. (2021). Exploring the toxic interactions between Bisphenol A and glutathione peroxidase 6 from Arabidopsis thaliana. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *259*, 119891. https://doi.org/https://doi.org/10.1016/j.saa.2021.119891

Xu, Z., Mahmood, K., & Rothstein, S. J. (2017). ROS Induces Anthocyanin Production Via Late Biosynthetic Genes and Anthocyanin Deficiency Confers the Hypersensitivity to ROS-Generating Stresses in Arabidopsis. *Plant and Cell Physiology*, *58*(8), 1364-1377. https://doi.org/10.1093/pcp/pcx073

Zi, X., Zhou, S., & Wu, B. (2022). Alpha-Linolenic Acid Mediates Diverse Drought Responses in Maize (Zea mays L.) at Seedling and Flowering Stages. *Molecules*, *27*(3). https://doi.org/10.3390/molecules27030771

Żmieńko, A., Samelak, A., Kozłowski, P., & Figlerowicz, M. (2014). Copy number polymorphism in plant genomes. *Theoretical and Applied Genetics*, *127*, 1-18.

Zubik-Duda, M., Luchowski, R., Maksim, M., Nosalewicz, A., Zgłobicki, P., Banaś, A. K., Grudzinski, W., & Gruszecki, W. I. (2023). The photoprotective dilemma of a chloroplast: To avoid high light or to quench the fire? *The Plant Journal*.