Mapping Host-Pathogen Interactions via Gene Co-expression Networks

Volkan Cevik
Biology and Biochemistry Department

Plant Pathology

Strategies of pathogenicity

A successful pathogen must:

- Find the host and attach to it

- Gain entry through the plant's impermeable defenses

- Avoid the plant's defense responses

- Grow and reproduce

- Spread to other plants

Rust uredospores *orientate germ tubes* on host surface to locate stomates for entry via appressoria

Types of Pathogens
Pathogens are Biotrophs, Necrotrophs or Hemibiotrophs

Arabidopsis thaliana

Necrotroph
Botrytis cinerea

Biotroph
Hyaloperonospora arabidopsidis

Hemibiotroph
Pseudomonas syringae

Necrotrophs secrete cell wall degrading enzymes and toxins killing cells and then consume the contents

Biotrophs live within host tissue without causing death
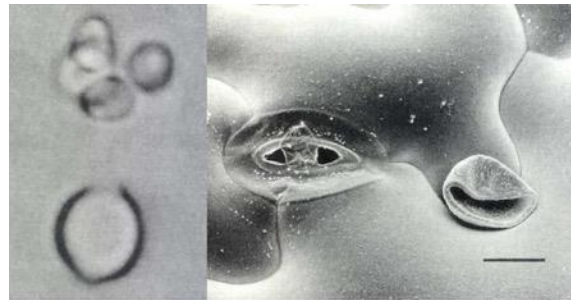
Hemibiotrophs can switch from biotroph to necrotroph

# *Albugo candida:* White Rust Disease of Brassica Crops

- Obligate biotrophic Oomycete
- Different races of *Albugo candida* can infect more than 200 plant species
- *Albugo candida* is one of the most important diseases of oilseed and vegetable Brassicas
- It has a strong immuno-suppression ability

Zoospores

Hyphae

Holub et al. 1995. MPMI 8: 916-928
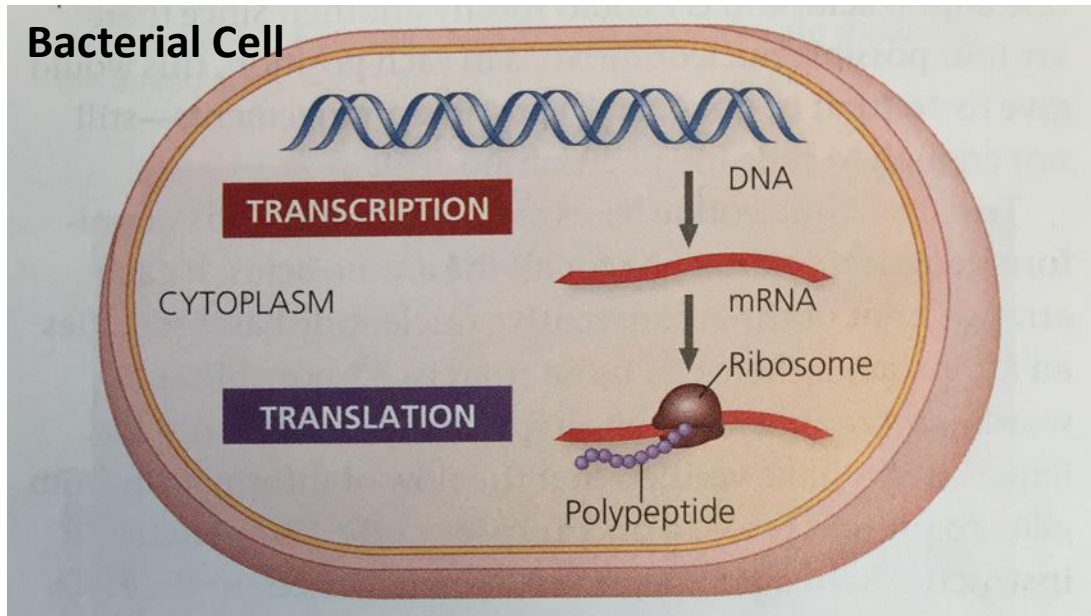
Holub et al. 1995. MPMI 8: 916-928
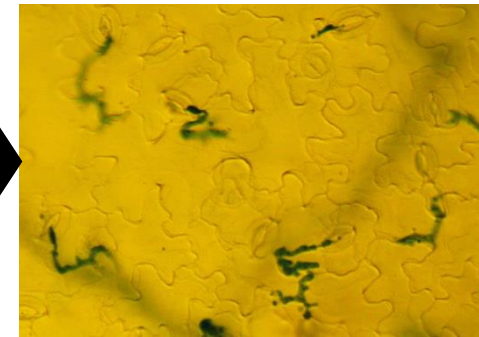
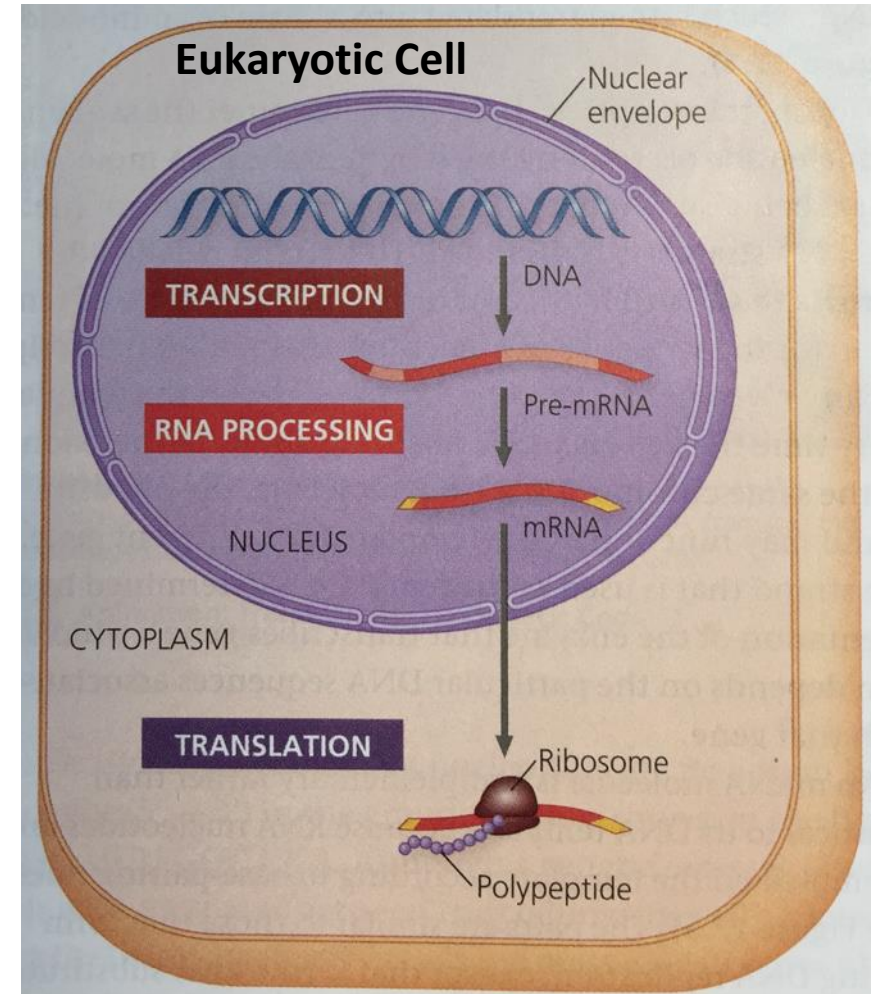**Asexual Life Cycle of *A. candida***

Pustules

Gene Expression

Genes (DNA) provide the instructions for making specific proteins. The bridge between DNA and protein synthesis is the nucleic acid RNA
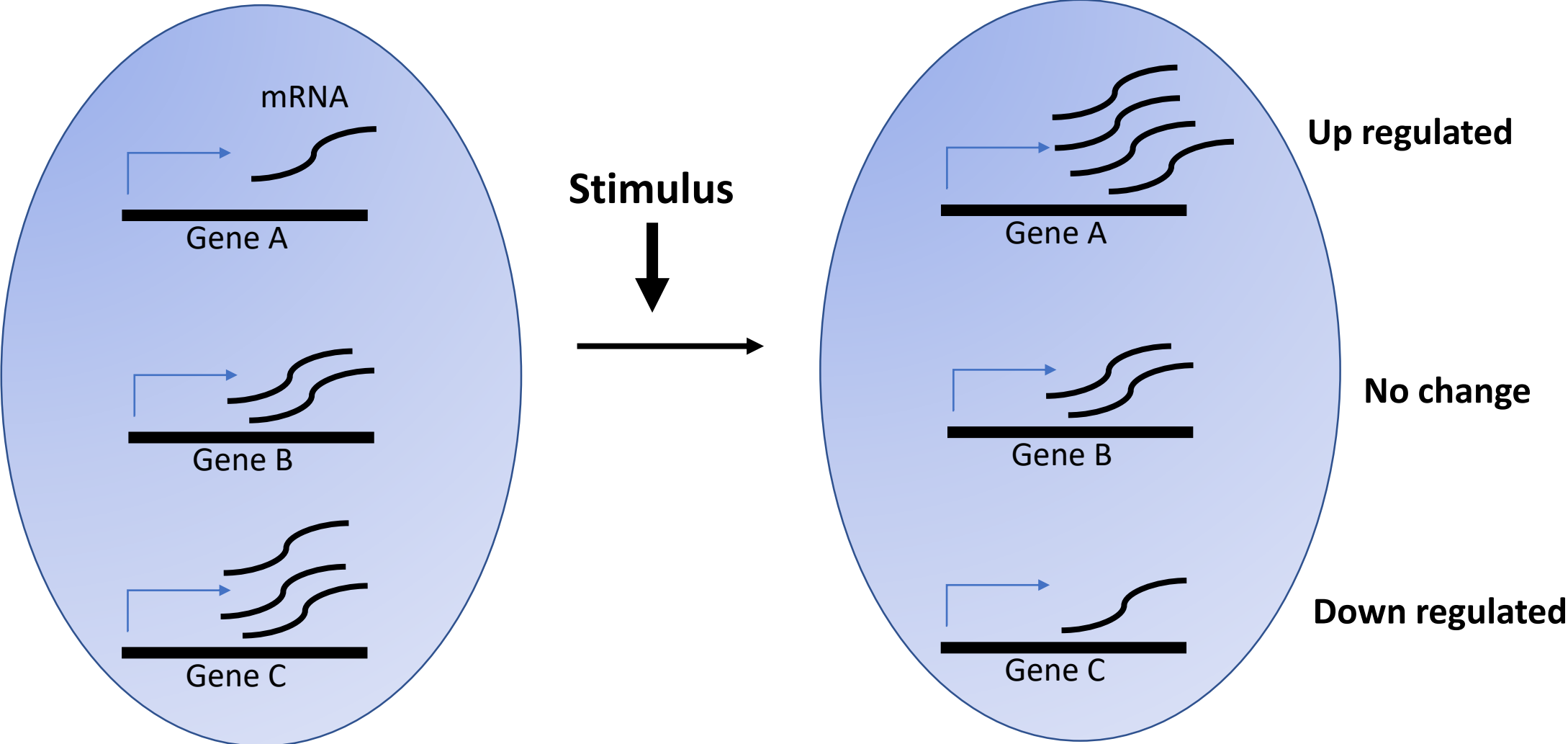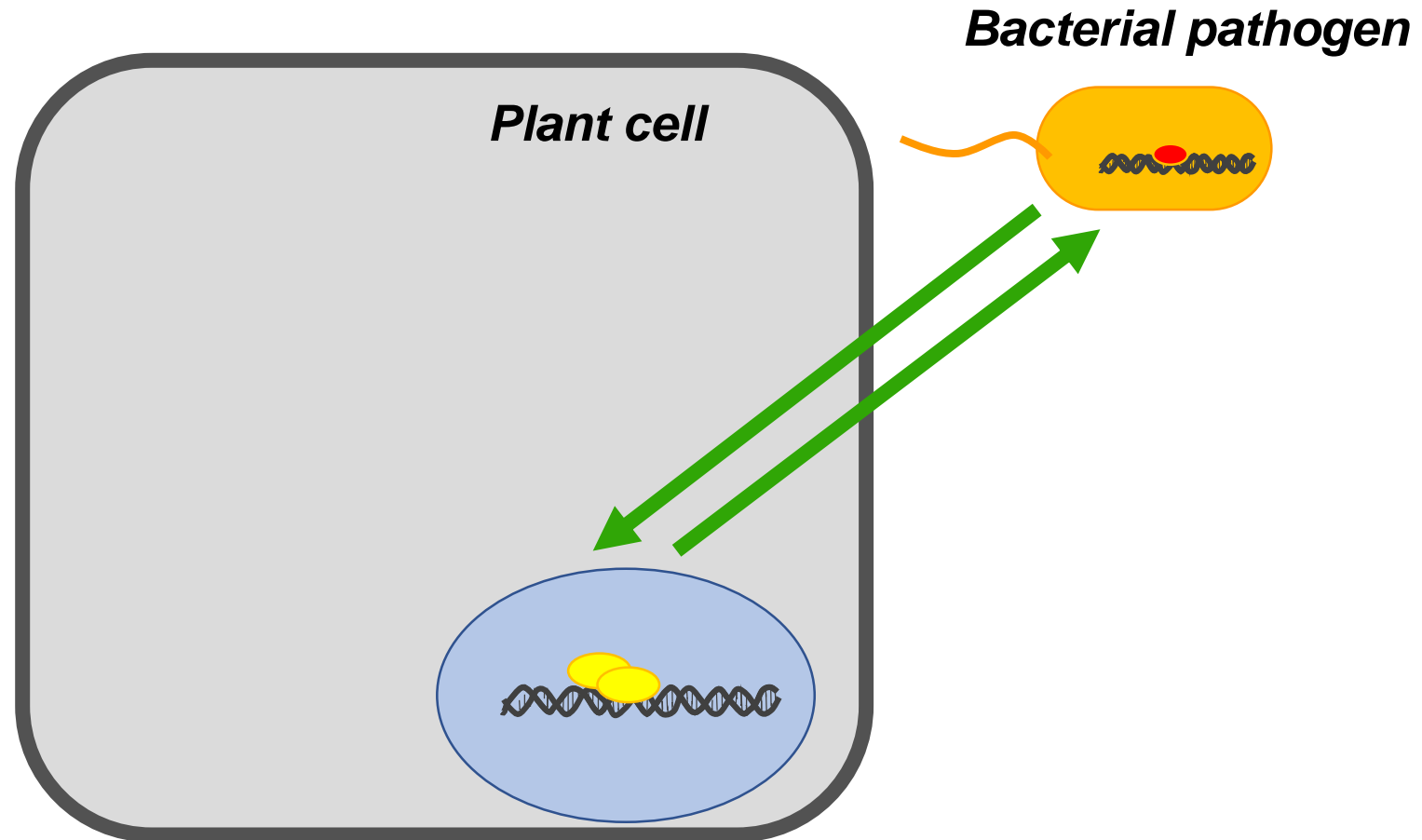
DNA > RNA > Protein

mRNA: Messenger RNA

Gene Expression

Gene Expression can change following external or internal stimuli

Plant-Pathogen Interactions involve large-scale transcriptional changes

Bacterial pathogen

Plant cell

How to Analyse Gene Expression
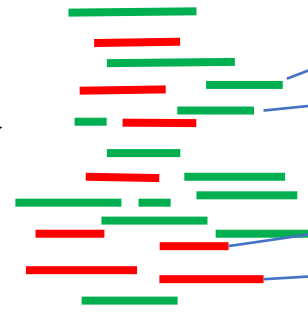## 'RNA Sequencing (RNA-Seq)'

Pathogen infected host

mRNA Extraction
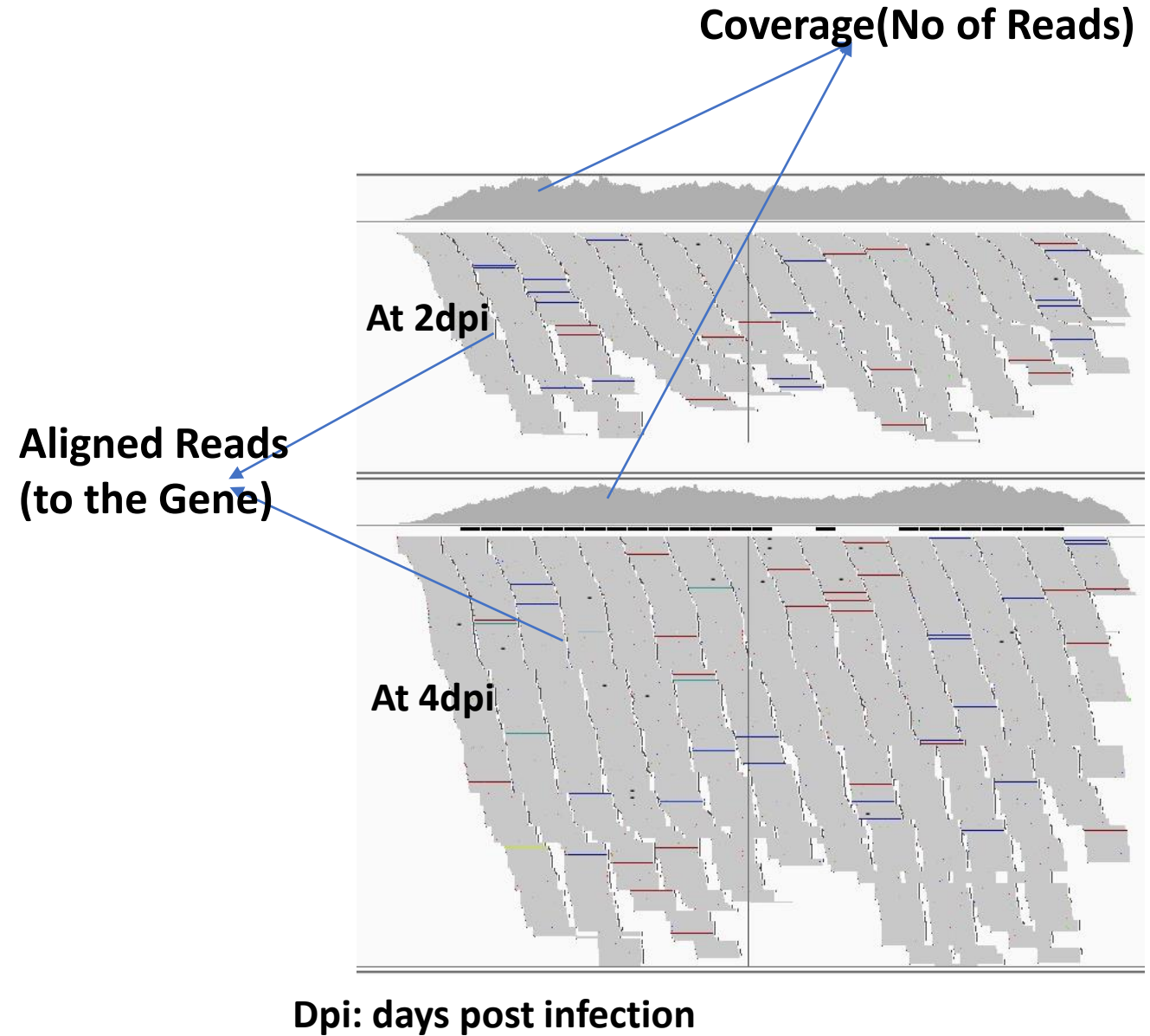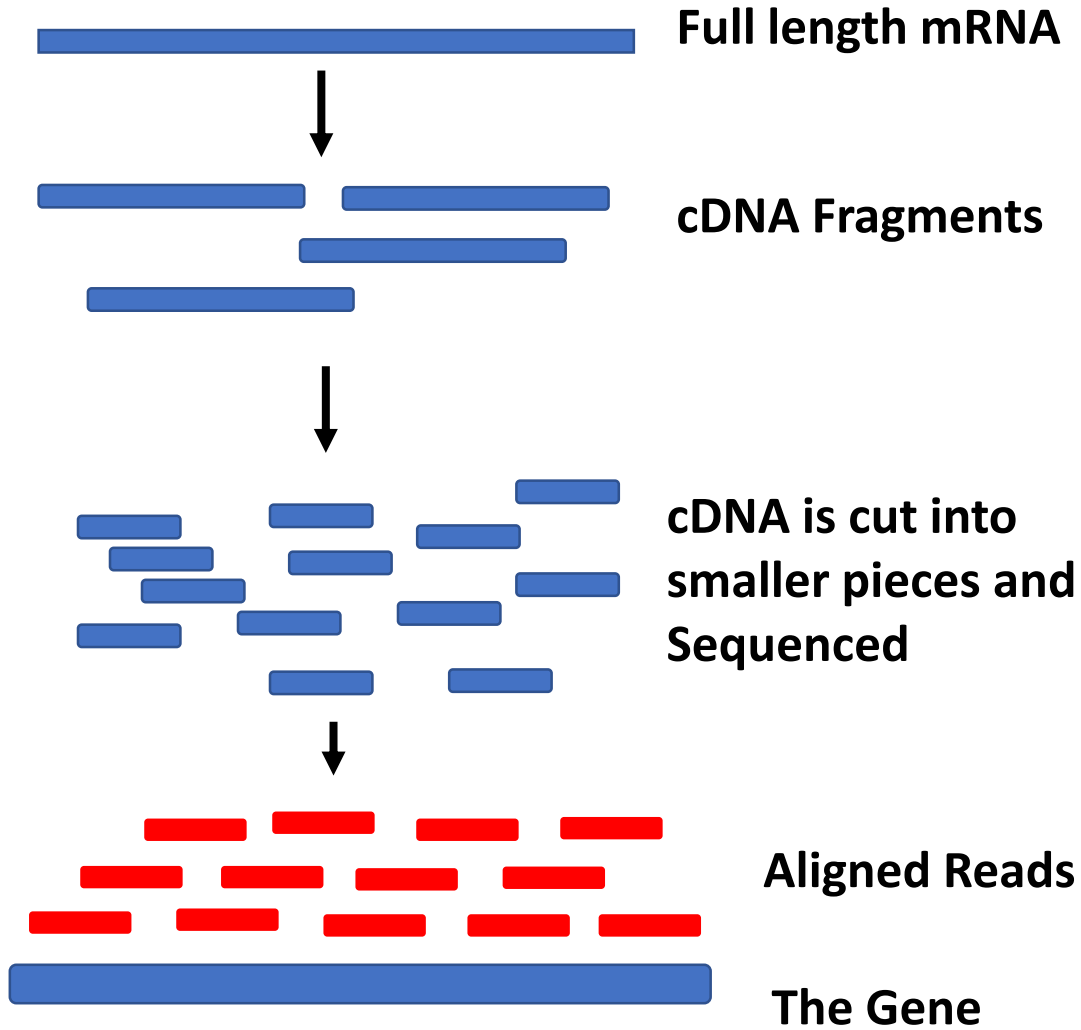
cDNA

Sequencing

AGCTAAGTCCCTC

GCTAAGTGATCGGG

TGGATACTTAGTC

CGTAAGTGATTTTAGC

Plant RNA

Pathogen RNA

Aligning short RNA-Seq reads

Full length mRNA

cDNA Fragments

cDNA is cut into smaller pieces and Sequenced

Aligned Reads

The Gene

Coverage(No of Reads)

Aligned Reads (to the Gene)

At 2dpi

At 4dpi

Dpi: days post infection

Bioinformatics Analysis of RNA-Seq

Counting the Aligned reads

number of reads/ fragments overlapping with the gene

↓

Normalization

RPKM: Reads/Fragments per
Kilobase of gene per Million reads
mapped

↓

TPM: Transcripts Per Million

$$TPM_i = \left(\frac{X_i}{l_i}\right) * \frac{1}{\sum_j \frac{X_j}{l_k}} * 10^6$$

gene read
counts per bp

all gene
counts per all
gene bp

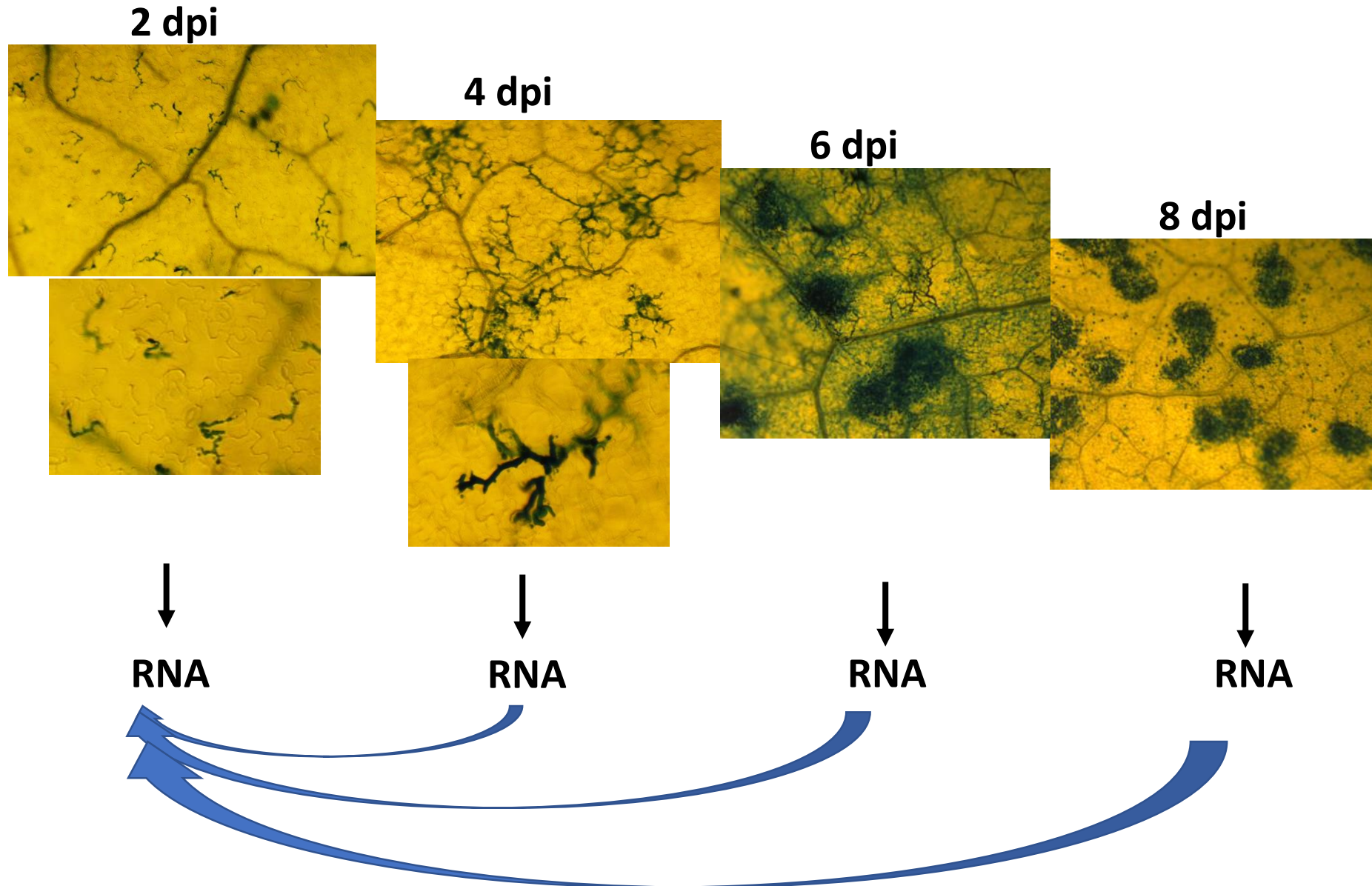**Differentially Expressed genes (Two-fold Down or Up regulated genes)**

Estimate magnitude of DE taking into account
differences in sequencing depth, technical, and
biological read count variability.

RNA-Seq Experiment during *Albugo candida - B. juncea (Oilseed mustard)* Interactions

2 dpi

4 dpi

6 dpi
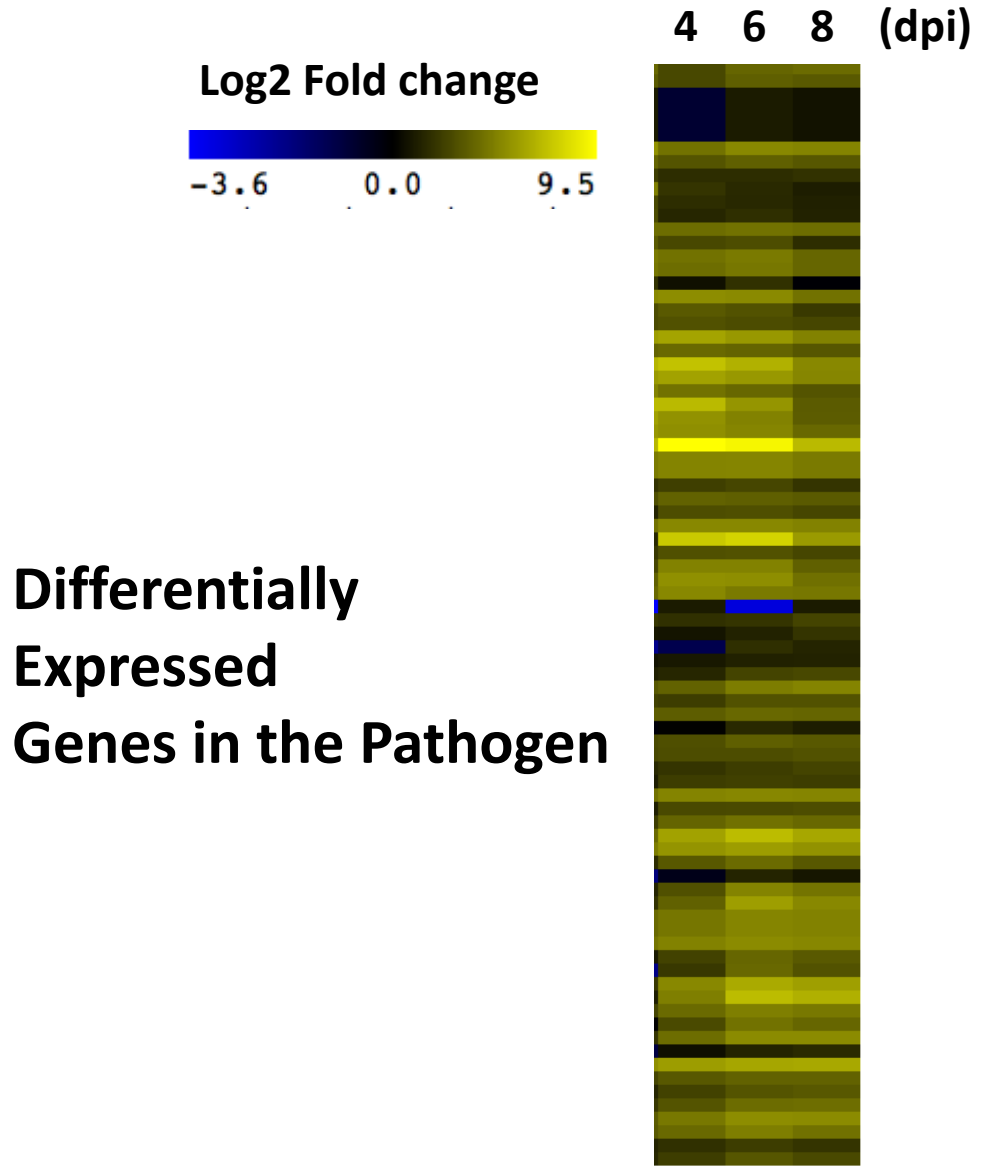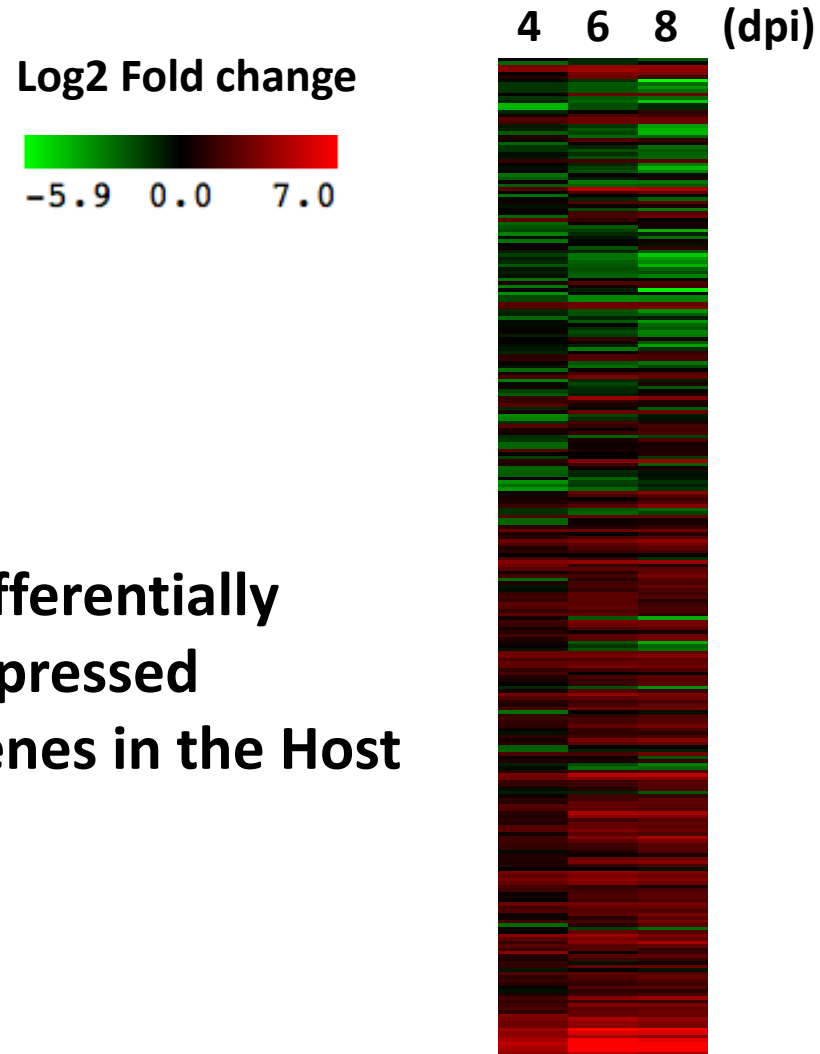
8 dpi

RNA

RNA

RNA

RNA
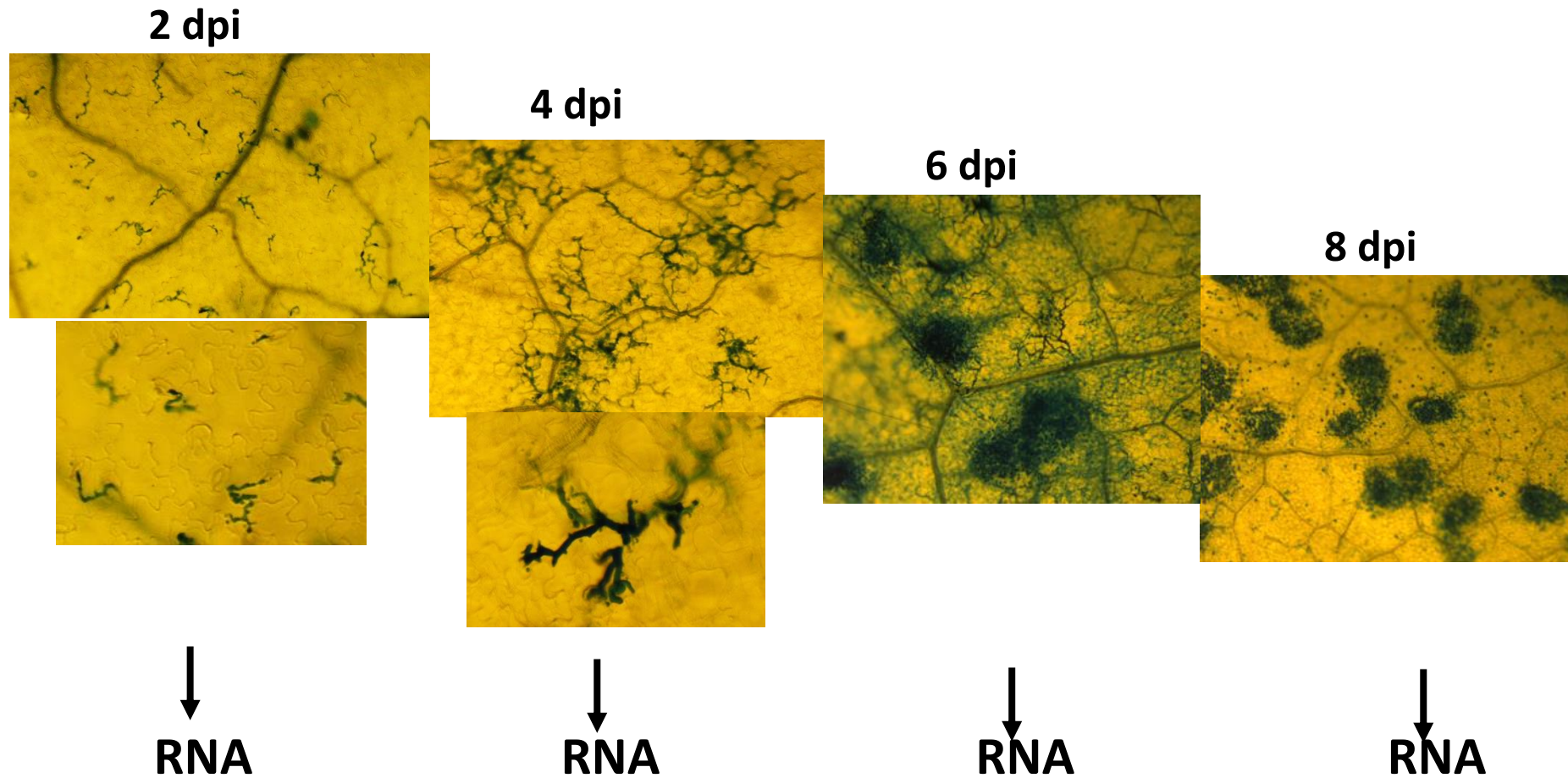
Three week old *B. juncea* plants inoculated with
Ac2v, stained with trypan blue
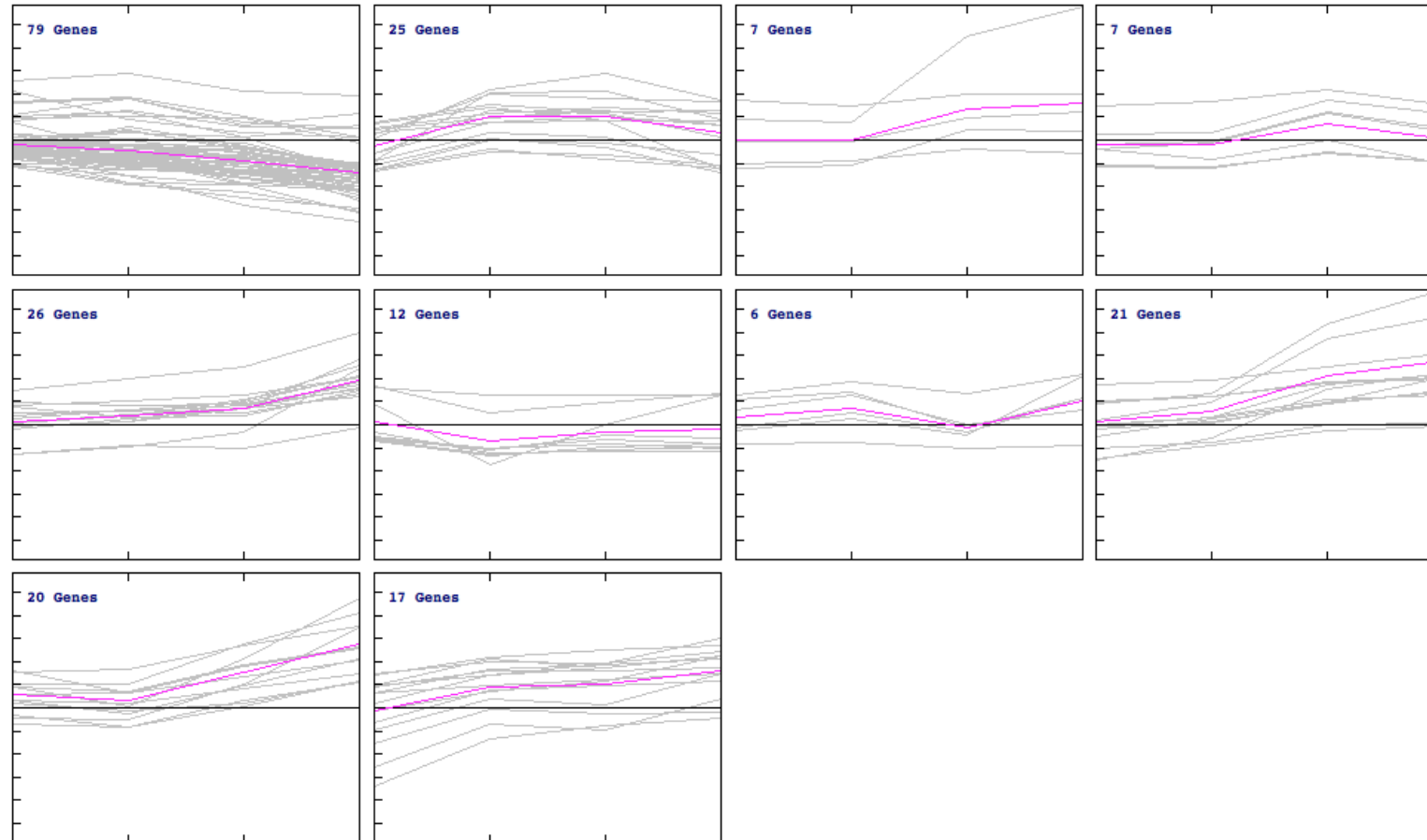Dpi: days post infection

Differential Gene Expression Analysis

Heat Map of Differentially expressed Genes in the pathogen and the host

K-means clustering (Using TPM) of the differentially expressed genes in the host at 2, 4, 6 and 8 dpi

The Challenge

Using the expression profiling of the differentially expressed genes both in the host and the pathogen, can we identify genes co-expression network

Normalized expression (TPM) of host/pathogen genes at different time points
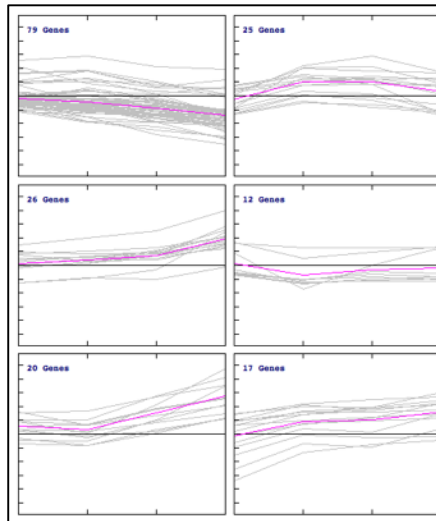
Gene co-expression network analyses

-Correlation Calculation (+ve/-ve) and Network Analysis

-Novel methods!