

Transcriptomics in Crop Research

Introduction to RNAseq technology

Mary-Ann Blätke
JJ Szymanski

Mon 6th

Gene to
transcript

Sequencing
technologies

Intro to
bash & setup

Tue 7th

Quantification

Getting data

Mapping

Wed 8th

Quality
check

Expression
units

Normalization

Fri 17th

Catching up

Q&A



...TGAGGTTGACCATTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTTAAGAAGATTCATGATTGTAGGTGAGGTTGACCATTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTTAAGA...

DNA

gene A

gene B

...TGAGGT**TGACCA**TTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTT**AAGAAG**ATTCA**TGATTGTAGG**TGAGGTTGACCATT**TGCCACACCTTGT**CCTTCCTCTGCTAGTCGTTTAAGA...

DNA

gene A

gene B

DNA

mRNA

transcription

AAACGGUGUGGAACAGGAAGGAGACGAUCAGCAA

ACUCCAACUGGUAACGGUGU

...TGAGGTGACCAATTGCCACACCTTGTCCTTCCTCTGCTAGTCGTTTAAGAAGATTCAATGATTGTAGGTGAGGTTGACCATTGCGCACACCTTGTCCTTCCTCTGCTAGTCGTTTAAGA...



gene A

gene B

DNA

mRNA

transcription

AAACGGUGUGGAACAGGAAGGAGACGAUCAGCAA

AAACGGUGUGGAACAGGAAGGAGACGAUCAGCAA

ACUCCAACUGGUAAACGGUGU

ACUCCAACUGGUAAACGGUGU

ACUCCAACUGGUAAACGGUGU

ACUCCAACUGGUAAACGGUGU

ACUCCAACUGGUAAACGGUGU

ACUCCAACUGGUAAACGGUGU

ACUCCAACUGGUAAACGGUGU

...TGAGGTGACCAATTGCCACACCTTGCTCCTTCCTCTGCTAGTCGTTTAAGAAGATTTCATGATTGTAGGTGAGGTTGACCATTGTCACACCTTGTCCTTCCTCTGCTAGTCGTTTAAGA...

gene A

gene B

DNA

cDNA

transcription & reverse transcription

TTTGCCACACCTTGCCTTCCTCTGCTAGTCGTTT

TTTGCCACACCTTGCCTTCCTCTGCTAGTCGTTT

TGAGGTTGACCATTGCGCACA

TGAGGTTGACCATTGCGCACA

TGAGGTTGACCATTGCGCACA

TGAGGTTGACCATTGCGCACA

TGAGGTTGACCATTGCGCACA

TGAGGTTGACCATTGCGCACA

TGAGGTTGACCATTGCGCACA

...TGAGGTTGACCAATTTGCCACACCTTGCCTTCCTCTGCTAGTCGTTTAAGAAGATTCAATGATTGTAGGTGAGGTTGACCATTGCGCACACCTTGTCTTCCTCTGCTAGTCGTTTAAGA...

gene A

gene B

DNA

cDNA

transcription & reverse transcription

TTT GCCACACCTTG TCCTTCCTC TGCTAGT CGTTT
TTTGCCAC ACCTTGTC TCCTCTGC TAGTCGTTT

TGAGGT TGACCATTT GCCACA
TGAGGTTGAC CATTTGCC ACA
TGAGG TTGACCATTG CCACA
TGA GGTGACCATTG CACA
TGAGGTT GACCAT TTGCCACA
TGAGGTTGA CCATTGCCA CA
TGAGGTTGACC ATTTGCCACA

...TGAGGT**TGACCA**TTTGCCACACCTTGTCCTTCCTCTGCTAGTCGTTT**AAGAAG**ATTCA**TGATTGTAGG**TGAGGTTGACCATT**TGCCACACCTTGT**CCTTCCTCTGCTAGTCGTTTAAGA...

gene A

gene B

DNA

cDNA

A bag of sequenced reads

TGAGGT TGACCATT GGCACA TGAGGTTGA CCATTGCCA CA
TGAGGTTGACC ATTTGCCACA
TGAGG TTGACCATTG CCACA TGAGGTTGAC CATTGCC ACA
TTTGCCAC ACCTTGTCC TTCCTCTGC TAGTCGTTT
TGAGGTT GACCAT TTGCCACA TGA GGTGACCATTG CACA
TTT GCCACACCTG TCCTTCCTC TGCTAGT CGTTT

gene A

gene B

...TGAGGT**TGACCA**TTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTT**AAGAAG**ATTCA**TGATTGTAGG**TGAGGTTGACCATT**TGCCACACCTTGT**CCTTCCTCTGCTAGTCGTTTAAGA...

DNA

cDNA

A library!

TGAGGT TGACCATT GGCACA TGAGGTTGA CCATTGCCA CA
TGAGGTTGACC ATTTGCCACA
TGAGG TTGACCATTG CCACA TGAGGTTGAC CATTGCC ACA
TTTGCCAC ACCTTGTCC TTCCTCTGC TAGTCGTTT
TGAGGTT GACCAT TTGCCACA TGA GGTGACCATTG CACA
TTT GCCACACCTTG TCCTTCCTC TGCTAGT CGTTT

gene A

gene B

DNA

cDNA

transcription & reverse transcription

A library!

TTGAGGT TGACCA TTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTT AAGAAGATTCA TGATTGTAGG TGAGGTTGACCATT TGCCACACCTTGT CCTTCCTCTGCTAGTCGTTT AAGA...

TTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTT

TGAGGTTGACCATTTGCCACA

TTGAGGT TGACCATTT GCCACA TGAGGTTGA CCATTGCCA CA
TGAGGTTGACC ATTTGCCACA
TGAGG TTGACCATTTG CCACA TGAGGTTGAC CATTGCCC ACA
TTTGCCAC ACCTTGTCC TTCCTCTGC TAGTCGTTT
TGAGGTT GACCAT TTGCCACA TGA GGTGACCATTGTC CACA
TTT GCCACACCTTG TCCTTCCTC TGCTAGT CGTTT

gene A

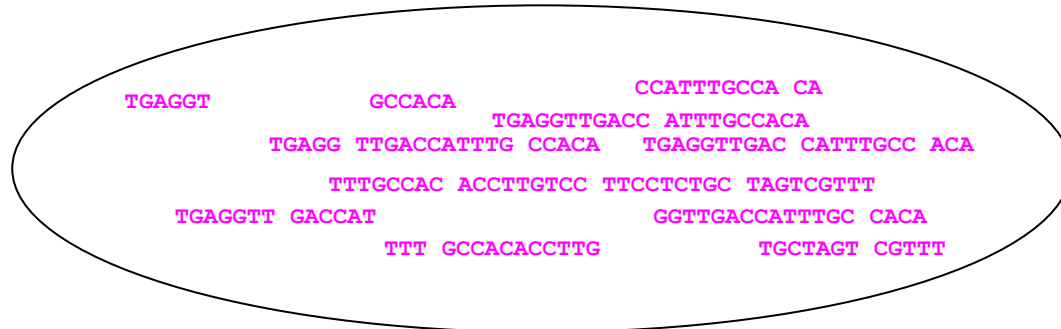
gene B

DNA

cDNA

transcription & reverse transcription

A library!



...TGAGGT TGACCA TTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTT AAGAAGATTCA TGATTGTAGG TGAGGTGACCATTTGCCACACCTTGT CCTTCCTCTGCTAGTCGTTT AAGA...

TTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTT
TCCTTCCTC

TGAGGTGACCATTTGCCACA
TGAGGTGA
TTGCCACA
TGA
TGACCATTT

gene A

gene B

DNA

cDNA

transcription & reverse transcription

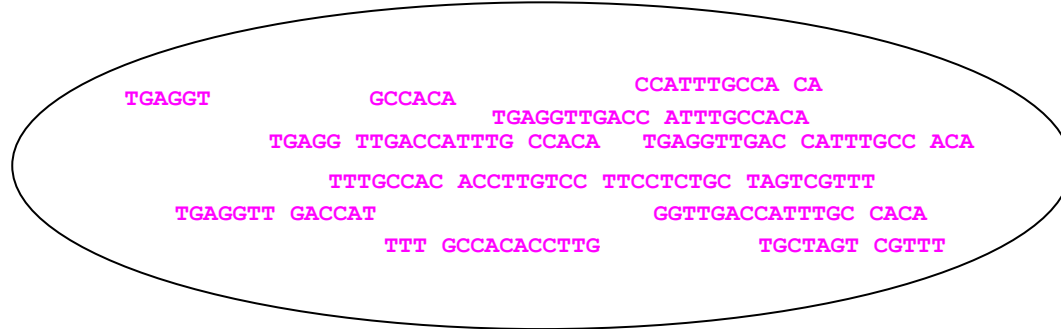
Mapping
the reads

A library!

...TGAGGTTGACCAATTGCCCACACCTTGTCTTCCTCTGCTAGTCGTTTAAGAAGATTTCATGATTGTAGGTGAGGTTGACCATTGCCCACACCTTGTCTTCCTCTGCTAGTCGTTTAAGA...

TTTGCCACACCTTGTCTTCCTCTGCTAGTCGTTT
TCCTTCCTC

TGAGGTTGACCATTGCCCACA
TGAGGTTGA
TGA
TTGCCACA
TGACCATT



What data do we need to start an RNAseq project?

gene A

gene B

DNA

cDNA

transcription & reverse transcription

A library

TTGAGGT TGACCAATT TGCCACACCTT GTCCTTCCTCTGCTAGTCGTTT AAGAAGATTCA TGATTGTAGG TGAGGTTGACCATT TGCCACACCTT GTCCTTCCTCTGCTAGTCGTTT AAGA...

TTTGCCACACCTT GTCCTTCCTCTGCTAGTCGTTT

TGAGGTTGACCATT TGCCACA

TTGAGGT TGACCAATT GGCACA TGAGGTTGA CCATTGCCA CA

TGAGGTTGACC ATTTGCCACA



TGAGG TTGACCATTG CCACA TGAGGTTGAC CATTGCCC ACA

TTTGCCAC ACCTTGTC TCCTCTGC TAGTCGTTT

TGAGGTT GACCAT TTGCCACA TGA GGTGACCATTG CACA

TTT GCCACACCTTG TCCTTCCTC TGCTAGT CGTTT


Get the reference genome

 **EnsemblPlants**  [HMMER](#) | [BLAST](#) | [BioMart](#) | [Tools](#) | [Downloads](#) | [Help & Docs](#) | [Blog](#)


Search: for


e.g. [Carboxy*](#) or [chx28](#)


All genomes





[View full list of all species](#)


Favourite genomes 


 **Arabidopsis thaliana**
TAIR10

 **Oryza sativa Japonica Group**
IRGSP-1.0

 **Triticum aestivum**
IWGSC

 **Hordeum vulgare**
MorexV3_pseudomolecules_assembly

 **Zea mays**
Zm-B73-REFERENCE-NAM-5.0

 **Physcomitrium patens**
Phypa_V3

Get the reference genome

e!EnsemblPlants | [HMMER](#) | [BLAST](#) | [BioMart](#) | [Tools](#) | [Downloads](#) | [Help & Docs](#) | [Blog](#)

Search: for







e.g. [Carboxy*](#) or [chx28](#)

All genomes

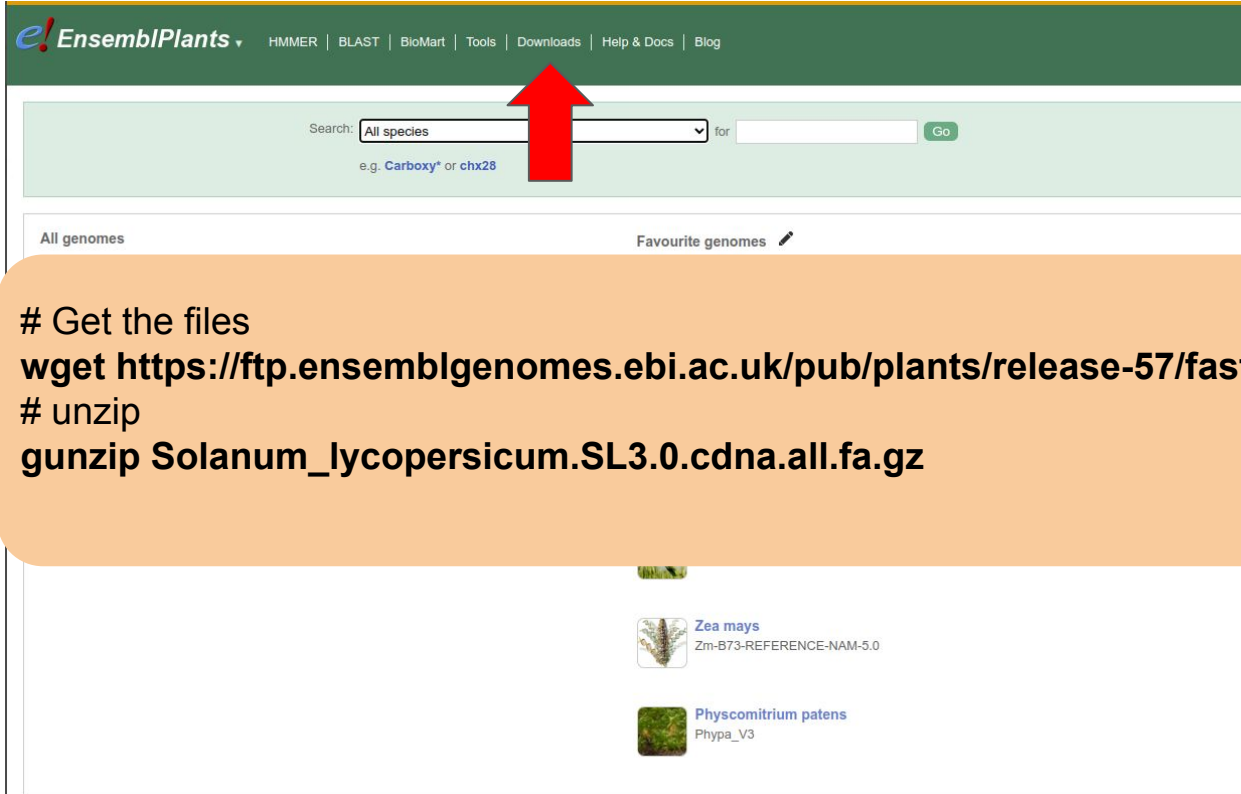
-- Select a species --

[View full list of all species](#)

Favourite genomes

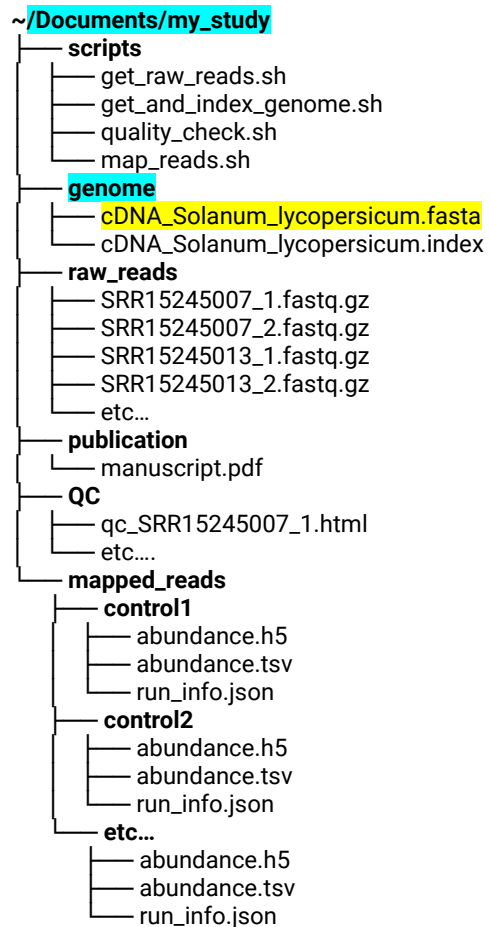
-  **Arabidopsis thaliana**
TAIR10
-  **Oryza sativa Japonica Group**
IRGSP-1.0
-  **Triticum aestivum**
IWGSC
-  **Hordeum vulgare**
MorexV3_pseudomolecules_assembly
-  **Zea mays**
Zm-B73-REFERENCE-NAM-5.0
-  **Physcomitrium patens**
Phypa_V3

Get the reference genome




The screenshot shows the Ensembl Plants website interface. At the top is a dark green navigation bar with the 'e!EnsemblPlants' logo and links for 'HMMER', 'BLAST', 'BioMart', 'Tools', 'Downloads', 'Help & Docs', and 'Blog'. Below this is a light green search bar with the text 'Search: All species' and a dropdown arrow. A large red arrow points to this dropdown menu. To the right of the search bar is a 'Go' button. Below the search bar, there are two tabs: 'All genomes' and 'Favourite genomes'. The 'All genomes' tab is active, showing a list of genomes. The first two visible entries are 'Zea mays' (Zm-B73-REFERENCE-NAM-5.0) and 'Physcomitrium patens' (Phypa_V3). Each entry has a small thumbnail image to its left.

Get the files
wget https://ftp.ensemblgenomes.ebi.ac.uk/pub/plants/release-57/fasta/solanum_lycopersicum/cdna/
unzip
gunzip Solanum_lycopersicum.SL3.0.cdna.all.fa.gz



From SRA search

 **National Library of Medicine**
National Center for Biotechnology Information

Log in


SRA

SRA

Search

Advanced

Help



SRA - Now available on the cloud

Sequence Read Archive (SRA) data, available through multiple cloud providers and NCBI servers, is the largest publicly available repository of high throughput sequencing data. The archive accepts data from all branches of life as well as metagenomic and environmental surveys. SRA stores raw sequencing data and alignment information to enhance reproducibility and facilitate new discoveries through data analysis.

Getting Started

[Documentation](#)

[How to submit](#)

[How to search and download](#)

[How to use SRA in the cloud](#)

[Submit to SRA](#)

Tools and Software

[Download SRA Toolkit](#)

[SRA Toolkit Documentation](#)

[SRA-BLAST](#)

[SRA Run Browser](#)

[SRA Run Selector](#)

Related Resources


[Submission Portal](#)


[dbGaP Home](#)


[BioProject](#)


[BioSample](#)


FOLLOW NCBI














Connect with NLM

National Library of Medicine
8600 Rockville Pike
Bethesda, MD 20894

Web Policies
FOIA
HHS Vulnerability Disclosure

Help
Accessibility
Careers

NLM | NIH | HHS | USA.gov

<https://www.ncbi.nlm.nih.gov/sra>

From SRA search

SRA Advanced Search Builder

((("solanum lycopersicum"[Organism]) AND "illumina"[Platform]) AND "strategy rna seq"[Properties])

[Edit](#)[Clear](#)

Builder

Organism

"solanum lycopersicum"[Organism]

⊖

[Show index list](#)

AND

Platform

"illumina"[Platform]

⊖

[Show index list](#)

AND

Properties

strategy rna seq[Properties]

strategy rip seq (20962)

strategy rna seq (4704915)

strategy selex (39382)

strategy ssrna seq (4703)

strategy synthetic long read (11812)

strategy targeted capture (391761)

strategy tethered chromatin conformation capture (654)

strategy tn seq (12547)

strategy validation (99)

strategy wgs (17036)

⬆

⬇

⬆

[Previous 200](#)

[Next 200](#)

[Refresh index](#)

AND

All Fields

⊖ ⊕

[Show index list](#)

Search

or [Add to history](#)

From SRA search

SRA Advanced Search Builder

((("solanum lycopersicum"[Organism]) AND "illumina"[Platform]) AND "strategy rna seq"[Properties]

[Edit](#)

[Clear](#)

Builder

Get the files

wget https://sra-downloadb.be-md.ncbi.nlm.nih.gov/sos3/sra-pub-zq-22/SRR010/056/SRR10056916.sralite.1

Convert SRA format to fastq.gz

fastq-dump --split-files --gzip -A SRR10056916 SRR10056916.sra

strategy validation (99)
strategy.wcs (17036)

[Refresh index](#)

AND ▼

All Fields ▼

Search

or [Add to history](#)

⌂ + [Show index list](#)

From ENA search



European Nucleotide Archive

Examples: [hs000006](#), [PRJNA562700](#), [PRJNA562700](#)

[Home](#) | [Submit](#) | [Search](#) | [Rulespace](#) | [About](#) | [Support](#)

We recommend that you subscribe to the ENA-announce mailing list for updates on ENA services.

Effective September 1st, 2023, our data retrieval APIs will implement enhanced performance measures. Each IP Address will be subject to a rate limit of 50 requests per second, ensuring optimized and efficient access to our APIs.

European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world's nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation. [More about ENA](#).

Access to ENA data is provided through the browser, through search tools, through large scale file download and through the API.



Latest ENA news

ENA: Improving spatio-temporal annotations Dec 1, 2021, 1:00:00 AM
The European Nucleotide Archive, along with its partners in the International Nucleotide Sequencing Consortium, has announced a new initiative to improve the accuracy and completeness of the ENA database. [Read more >](#)

Retirement of old ENA Browser on 5th August 2020 Jul 16, 2020, 2:00:00 AM
The new ENA Browser (<https://www.ebi.ac.uk/ena/browser/home>) has been running in parallel to our old Browser (<https://www.ebi.ac.uk/ena>) since mid 2019. [Read more >](#)

[See all news](#)



Please take our brief (9 minutes) survey about how ENA data is used : <https://forms.gle/hZRjcpEnGhAaVKGA>

1 Data Type

2 Query

3 Inclusion/Exclusion

4 Fields

5 Data Filters

Taxonomy and related

Geographical location

Geography

Collection event information

Sampling information

Sample state and conditions

Host information

Methodology

Sequencing information

Database record

File information

Accessions

Titles, aliases and descriptions

Sequenced

AND OR

NCBI Taxonomy

=

solanum lycopersicum

X

☒ NCBI taxonomic classification

Solanum lycopersicum L., 1753

4081

Solanum lycopersicum var. cerasiforme (Alef.) Fosberg, 1955

195583

Solanum lycopersicum var. cerasiforme x Solanum pimpinellifolium

2928697

Solanum lycopersicum x Solanum habrochaites

357543

Solanum lycopersicum x Solanum pennellii

672925

Solanum lycopersicum x Solanum peruvianum

95449

Solanum lycopersicum x Solanum pimpinellifolium

286530

Solanum lycopersicum x Solanum tuberosum

298315

Solanum



The European Nucleotide Archive (ENA) is part of the ELIXIR infrastructure
The ENA is an ELIXIR Core Data Resource. [Learn more >](#)



The European Nucleotide Archive (ENA) is a Global Core Biodata Resource
The ENA is a GBC Global Core Biodata Resource. [Learn more >](#)

EMBL-EBI is the home for big data in biology.

We help scientists exploit complex information to make discoveries that benefit humankind.

Source data of a manuscript



AoB PLANTS 2020, Vol. 12, No. 5

doi:10.1093/aobpla/plaa041

Advance Access Publication August 19, 2020
Studies

STUDIES

Transcriptome analysis in osmo-primed tomato seeds with enhanced longevity by heat shock treatment

Thiago Barbosa Batista¹, Geysson Javier Fernandez², Tiago Alexandre da Silva¹,
Júlio Maia¹ and Edvaldo Aparecido Amaral da Silva¹

¹Department of Plant Production, Sao Paulo State University (UNESP), Botucatu, Sao Paulo, Brazil, ²Institute of Biology, Antioquia University, Medellín, Antioquia, Colombia

*Corresponding author's e-mail address: amaral.silva@unesp.br

Associate Editor: Gabriela Augé

Form & Function, Chief Editor: Kate McCulloh

Abstract

Seed priming is widely used in commercial seeds and its main function is to accelerate and synchronize seed germination. Undesirably, primed seeds show reduced longevity and treatments like heat shock have been shown to improve longevity in primed seeds. Nonetheless, the effect of heat shock treatment on primed seeds at the mRNA level is not known. Thus, the aim of this work was to investigate the effect of heat shock treatment on the longevity of primed tomato (*Solanum lycopersicum*) seeds at the physiological and transcriptome levels. Tomato seeds were primed and dried (control). Alternatively, primed seeds were subjected to heat shock treatment (38 °C/32 % relative humidity) before drying. Germination, vigor and longevity were evaluated. Transcriptome analysis was performed by RNA sequencing (RNA-seq) from biological samples collected immediately after priming and another samples collected from primed seeds followed by the heat shock treatments. The gene expression was validated by quantitative real time PCR (RT-qPCR). We showed that applying heat shock treatment after priming increased germination speed, enhanced seed longevity and preserved the vigor during storage of primed tomato seeds. Through transcriptome analysis, 368 differentially expressed genes were identified, from which 298 genes were up-regulated and 70 were down-regulated. We showed the increase of mRNA levels of HEAT SHOCK FACTOR-like and HEAT SHOCK PROTEIN-like chaperone genes, suggesting the involvement of the proteins coded by these transcripts in the enhancement of longevity in primed tomato seeds. The heat shock treatment after priming enhances and preserves the vigor of tomato primed seeds during storage. In addition, improves seed longevity through the increase in the expression of transcripts related to protection by response to stress.

Keywords: Chaperone molecules; improved longevity; primed seed; seed conservation; seed quality; *Solanum lycopersicum* L.; storage.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7494243/>

Source data of a manuscript



AoB PLANTS 2020, Vol. 12, No. 5

doi:10.1093/aobpla/plaa041
Advance Access Publication August 19, 2020
Studies

STUDIES

Transcriptome analysis in osmo-primed tomato seeds with enhanced longevity by heat shock treatment

Thiago Barbosa Batista¹, Geysson Javier Fernandez², Tiago Alexandre da Silva¹, Júlio Maia¹ and Edvaldo Aparecido Amaral da Silva¹

¹Department of Plant Production, Sao Paulo State University (UNESP), Botucatu, Sao Paulo, Brazil, ²Institute of Biology, Antioquia University, Medellín, Antioquia, Colombia

*Corresponding author's e-mail address: amaral.silva@unesp.br

Associate Editor: Gabriela Augé

Form & Function. Chief Editor: Kate McCulloh

Abstract

Seed priming is widely used in commercial seeds and its main function is to accelerate and synchronize seed germination. Undesirably, primed seeds show reduced longevity and treatments like heat shock have been shown to improve longevity in primed seeds. Nonetheless, the effect of heat shock treatment on primed seeds at the mRNA level is not known. Thus, the aim of this work was to investigate the effect of heat shock treatment on the longevity of primed tomato (*Solanum lycopersicum*) seeds at the physiological and transcriptome levels. Tomato seeds were primed and dried (control). Alternatively, primed seeds were subjected to heat shock treatment (38 °C/32 % relative humidity) before drying. Germination, vigor and longevity were evaluated. Transcriptome analysis was performed by RNA sequencing (RNA-seq) from biological samples collected immediately after priming and another samples collected from primed seeds followed by the heat shock treatments. The gene expression was validated by quantitative real time PCR (RT-qPCR). We showed that applying heat shock treatment after priming increased germination speed, enhanced seed longevity and preserved the vigor during storage of primed tomato seeds. Through transcriptome analysis, 368 differentially expressed genes were identified, from which 298 genes were up-regulated and 70 were down-regulated. We showed the increase of mRNA levels of HEAT SHOCK FACTOR-like and HEAT SHOCK PROTEIN-like chaperone genes, suggesting the involvement of the proteins coded by these transcripts in the enhancement of longevity in primed tomato seeds. The heat shock treatment after priming enhances and preserves the vigor of tomato primed seeds during storage. In addition, improves seed longevity through the increase in the expression of transcripts related to protection by response to stress.

Keywords: Chaperone molecules; improved longevity; primed seed; seed conservation; seed quality; *Solanum lycopersicum* L.; storage.

Physiological assays

Seed germination and vigor. Four replications of 50 seeds were germinated in 9 cm Petri dishes with substrate of paper towel moistened with distilled water equivalent to 2.5 times its weight, at 25 °C, under 8 h of light and 16 h in the dark. The length of the primary root, ≥ 2 mm was used as the germination criterion. Data collection was done in different times after sowing; and ended when the germination rate reached 100 % or at 14 days. Seed vigor was determined by the calculation of the time to 50 % of germination (t50) through the analysis of cumulative germination data using the curve fitting module of the Germinator software package (Joosen *et al.* 2010).

Longevity. We used ageing protocol to assess seed longevity in which the seeds were placed in a support over a saturated solution of NaCl (75 % RH) at 35 °C in glass bottles hermetically sealed. During storage, the water content of *S. lycopersicum* seeds stabilized at 0.10 ± 0.007 g H₂O/g DW⁻¹, corresponding to ± 9.5 % on wet basis. At different time spans, seeds were imbibed and viability was assessed using the germination assay as described earlier. The different time spans were carried out considering the viability loss behaviour of each treatment group during storage. The viability data were transformed into probit to

libraries were 100 base pair (bp) paired-end sequenced. The data output in fastq file format contained sequence information, including the sequencing quality (Phred quality score). Average Phred scores of ≥ 20 per position were used for the alignment.

Read alignment and differentially expressed genes. Paired-end reads for mRNA were mapped to the *Solanum lycopersicum* release 39 reference genome using the default parameters of TopHat2 (Kim *et al.* 2013). Counts for RefSeq genes were obtained using HTSeq (Anders *et al.* 2015) and DESeq2 (Love *et al.* 2014) was used to normalize expression counts. The changes in gene expression were considered statistically significant when fold change ≥ 2 and P-values ≤ 0.05 . The RNAseq data was deposited in NCBI (BioProject PRJNA562700: <https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP220280>).

The analysis of principal components was made using all the genes expressed on the RNA seq data. The normalized count per gene was used and transformed to Z-score. This matrix was used to perform the PCA. For plotting the PCA results, we used the principal component one and two. The heatmap was generated using the normalized counts of the differentially expressed genes. Then we transformed it to z-score and plotted it using the package pheatmap of R.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7494243/>

From ENA search

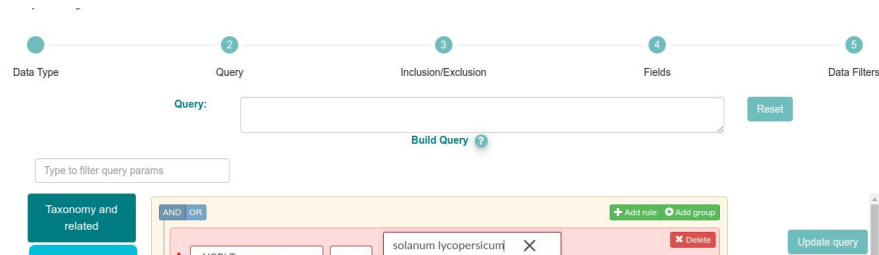


We recommend that you subscribe to the ENA-announce mailing list for updates on ENA services.

Effective September 1st, 2023, our data retrieval APIs will implement enhanced performance measures. Each IP Address will be subject to a rate limit of 50 requests per second, ensuring optimized and efficient access to our APIs.

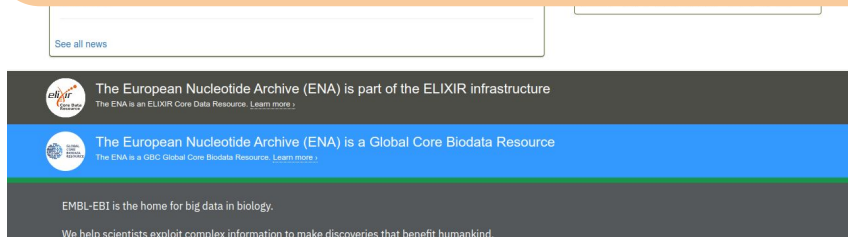
European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world's nucleotide sequencing information, covering raw sequencing data, sequence assembly information and



Get the files

```
wget -nc ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR152/098/SRR15244998/SRR15244998_1.fastq.gz -P fastq_files
```



Solanum lycopersicum x Solanum peruvianum 95449
Solanum lycopersicum x Solanum pimpinellifolium 284530
Solanum lycopersicum x Solanum tuberosum 298315
Solanum

~/Documents/my_study

scripts

- get_raw_reads.sh
- get_and_index_genome.sh
- quality_check.sh
- map_reads.sh

genome

- cDNA_Solanum_lycopersicum.fasta
- cDNA_Solanum_lycopersicum.index

raw_reads

- SRR15245007_1.fastq.gz
- SRR15245007_2.fastq.gz
- SRR15245013_1.fastq.gz
- SRR15245013_2.fastq.gz
- etc...

publication

- manuscript.pdf

QC

- qc_SRR15245007_1.html
- etc....

mapped_reads

control1

- abundance.h5
- abundance.tsv
- run_info.json

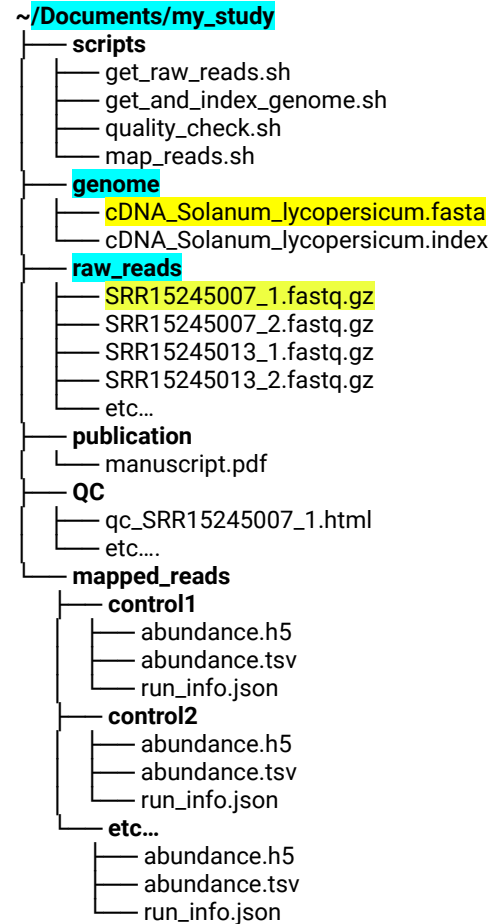
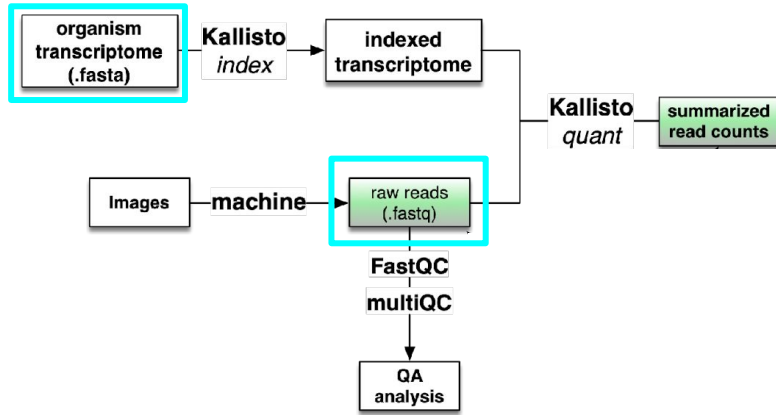
control2

- abundance.h5
- abundance.tsv
- run_info.json

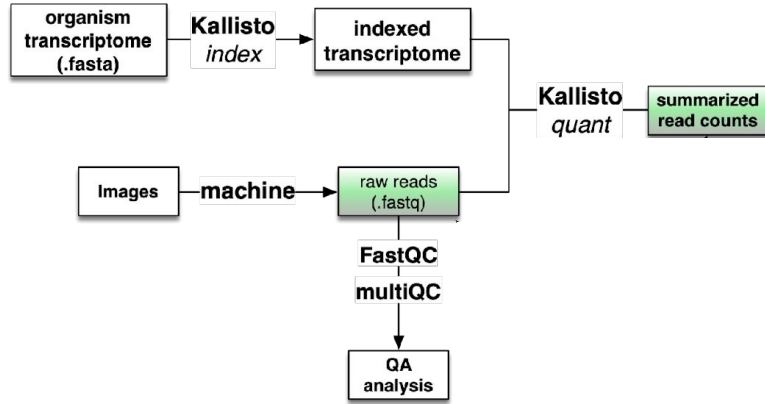
etc...

- abundance.h5
- abundance.tsv
- run_info.json

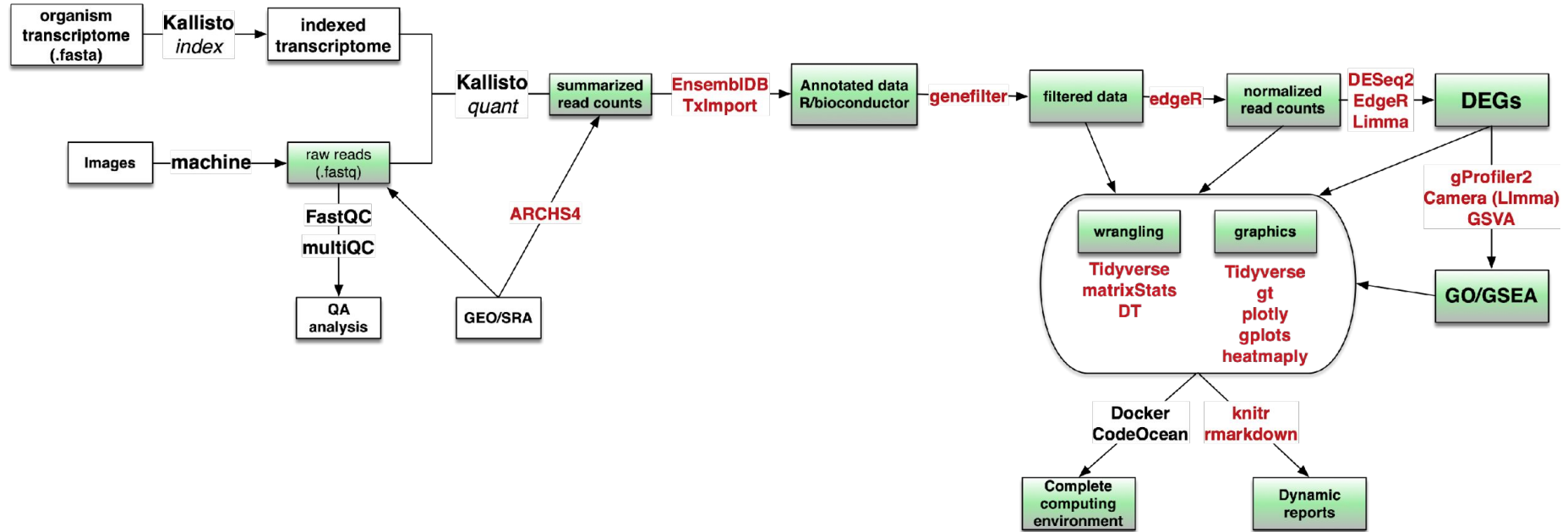
RNAseq analysis pipeline



The RNAseq analysis pipeline



The RNAseq analysis pipeline



The End