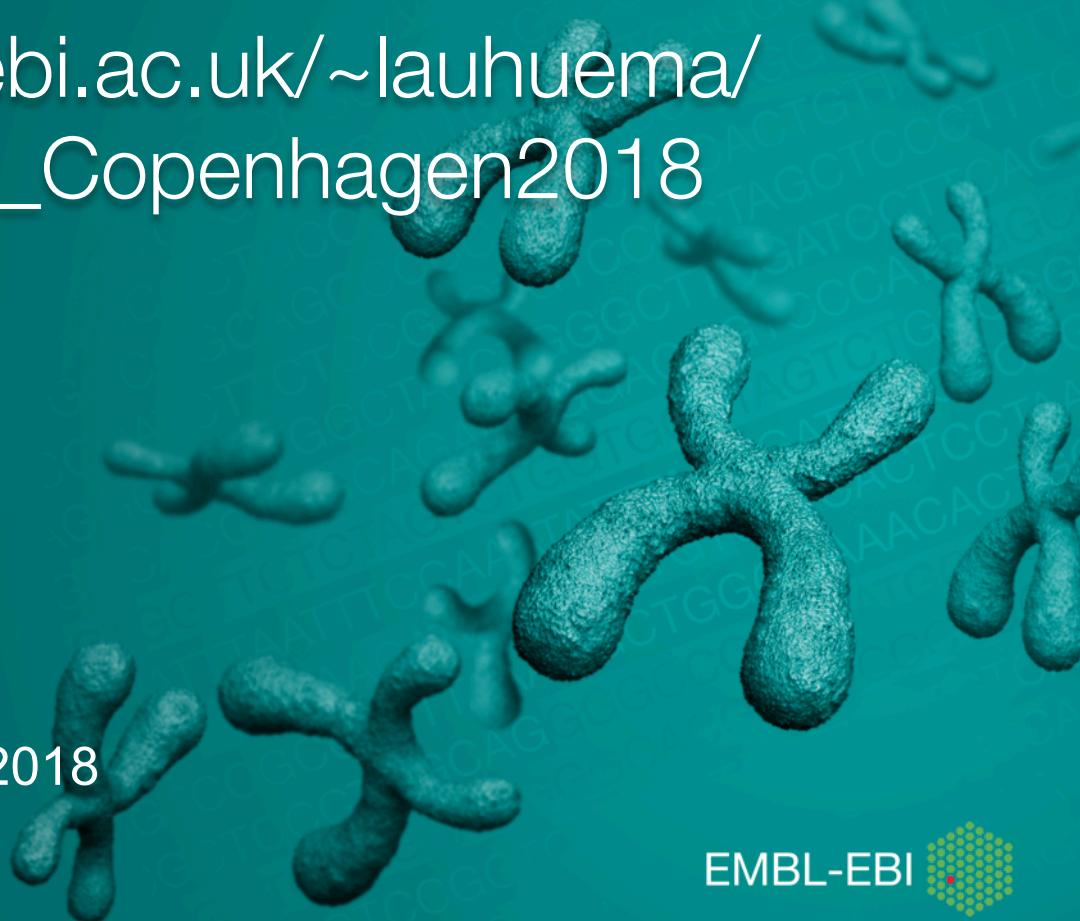


EMBL-EBI workshop: Exploring plant gene expression in Expression Atlas

[https://www.ebi.ac.uk/~lauhuema/
workshop/fg_Copenhagen2018](https://www.ebi.ac.uk/~lauhuema/workshop/fg_Copenhagen2018)

Laura Huerta, PhD
Senior Scientific Curator
lauhuema@ebi.ac.uk
Copenhagen, 10 September 2018





UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

A bit about your speaker...



- *Agronomic Engineer, major in Biotechnology*

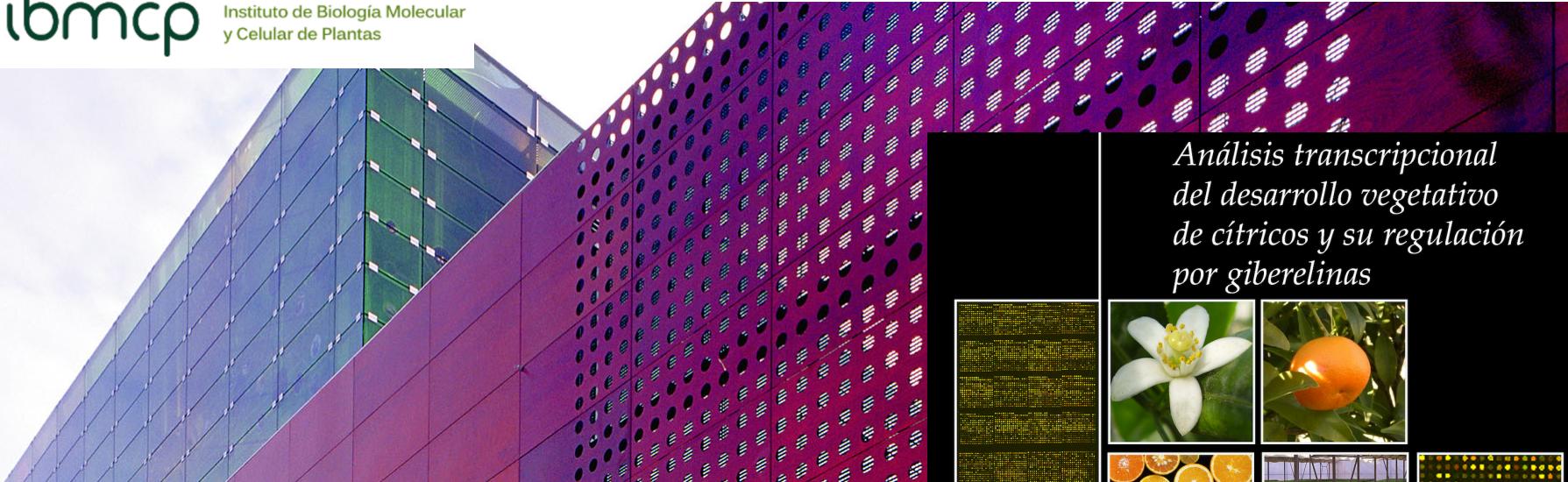
- *Cloning and characterisation of a GA 20-oxidase gene from Citrus sinensis (L.) Osbeck.*

Prof. José Luis García Martínez (CSIC)

A bit about your speaker...

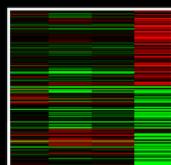
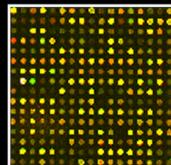
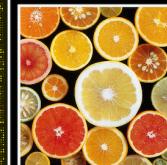


Instituto de Biología Molecular
y Celular de Plantas



<http://www.ibmcp.csic.es>

*Análisis transcripcional
del desarrollo vegetativo
de cítricos y su regulación
por giberelinas*



TESIS DOCTORAL
Laura Huerta Martínez

A bit about your speaker...

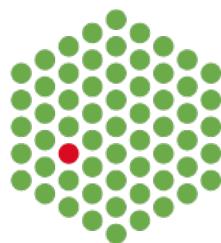




WELLCOME
GENOME
CAMPUS



EMBL-EBI



EMBL-EBI workshop: Exploring plant gene expression in Expression Atlas

Introduction to EMBL-EBI resources

Laura Huerta, PhD

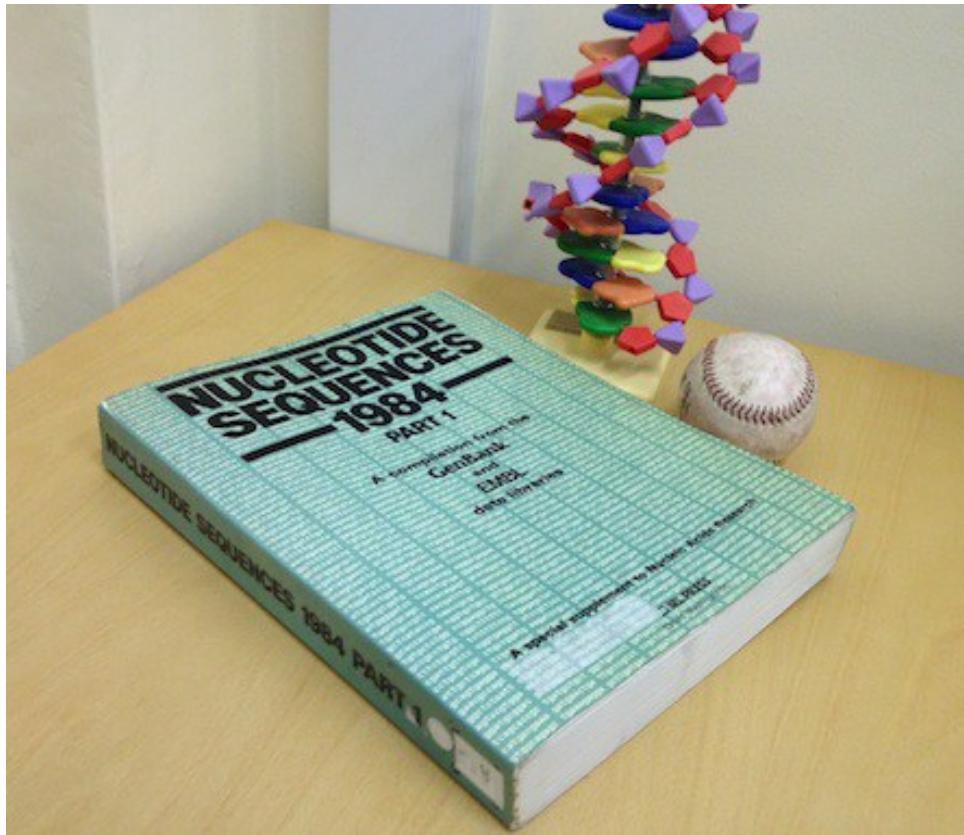
Senior Scientific Curator

lauhuema@ebi.ac.uk

Copenhagen, 10 September 2018



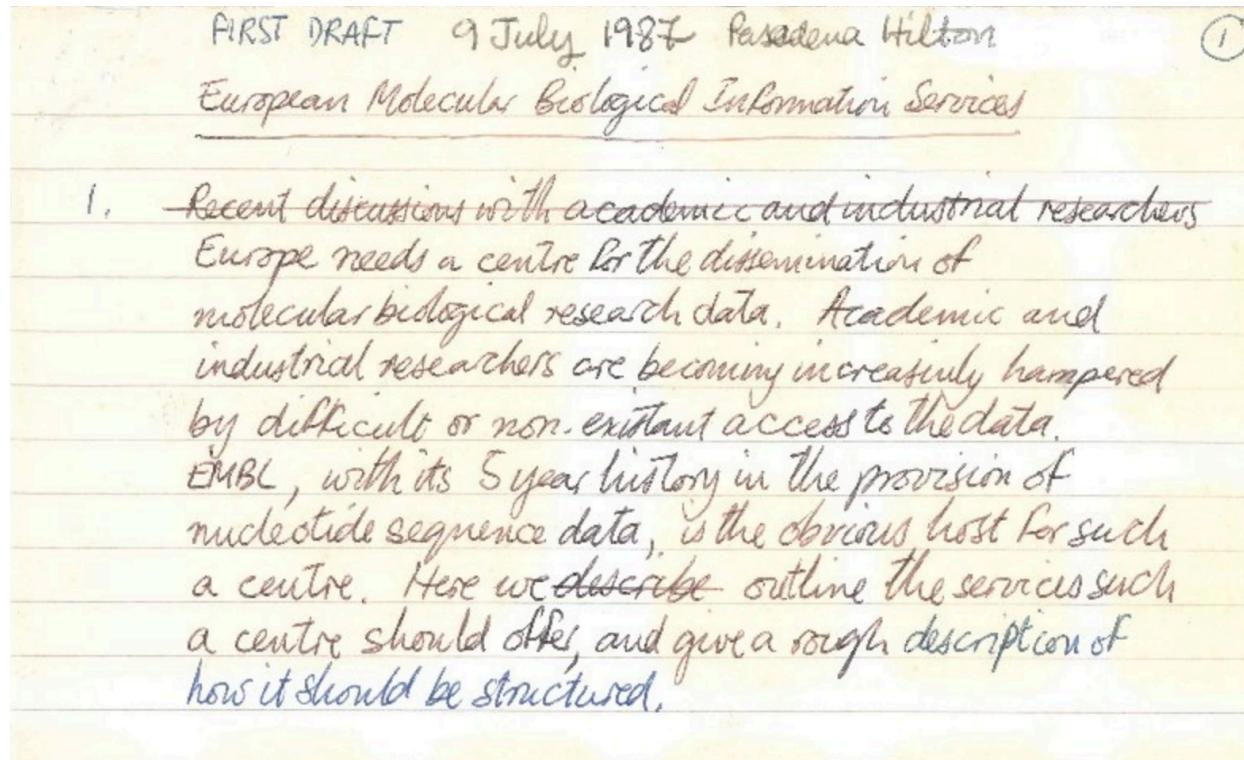
Life science: a revolution in biological data



Nucleotide sequences 1984 Part 1 A compilation
from the GenBank and EMBL data libraries

Life science: a revolution in biological data

The first draft of Graham Cameron's proposal for a “European Molecular Biological Information Services”.



Life science: a revolution in biological data



Life science: a revolution in biological data



The European Bioinformatics Institute (EBI)

The home for big data in biology



To provide freely available data and bioinformatics services to the scientific community

www.ebi.ac.uk/

What is EMBL-EBI?

- Europe's home for biological data services, research and training
- A trusted data provider for the life sciences:
 - 150 Petabytes of storage
 - > 40,000 CPU Cores
- Part of the European Molecular Biology Laboratory, an intergovernmental research organisation
- International: 600 members of staff from 60 nations
- Home of the ELIXIR Technical hub.

The European Molecular Biology Laboratory

+60 nationalities

Heidelberg, Germany



Main Laboratory

>1800 personnel

Hinxton, Cambridge, UK



Bioinformatics

6 sites in Europe

Grenoble, France



Structural Biology



Tissue Biology, Disease Modeling
Barcelona, Spain



Mouse Biology
Monterotondo, Rome, Italy



Structural Biology
Hamburg, Germany

EMBL member states



Austria, Belgium, Croatia, Czech Republic, **Denmark**, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Luxembourg, Malta, Montenegro, the Netherlands, Norway, Portugal, Slovakia, Republic, Spain, Sweden, Switzerland and the United Kingdom

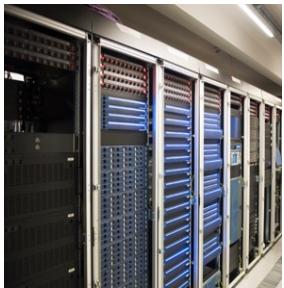
Associate member states: Argentina, Australia

Prospect member states: Lithuania, Poland



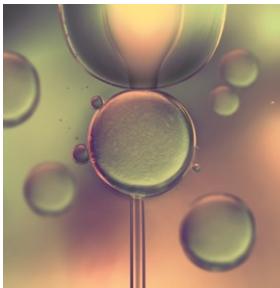
What is EMBL-EBI? – Our mission

Services



Deliver scientific services

Research



Deliver excellent research

Training



Train the next generation of scientists

Industry



Engage with European industry

ELIXIR



Coordinate bioinformatics and data in Europe

Bioinformatics services

Data and tools to support life science
research

www.ebi.ac.uk/services

What services do we provide?

...provide tools to help researchers use it

Labs around the world send us their data and we...

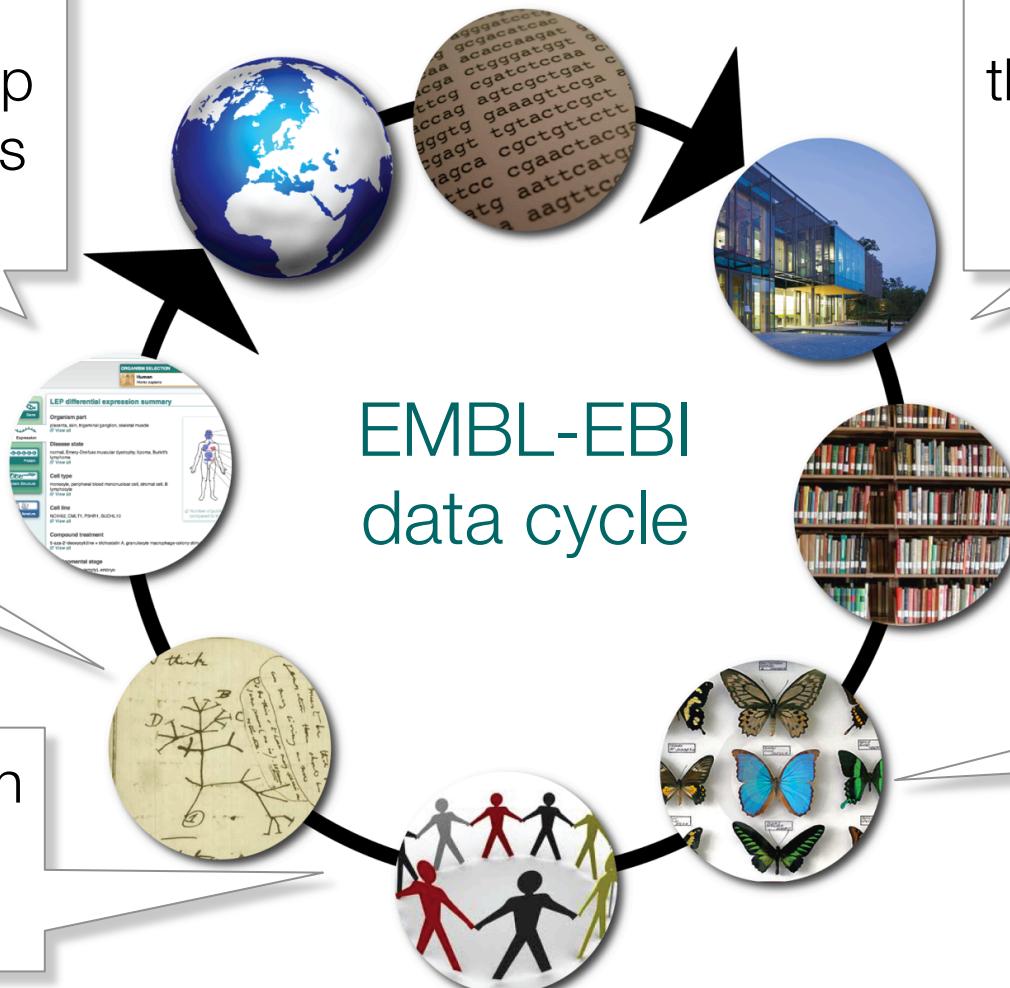
EMBL-EBI data cycle

Analyse it

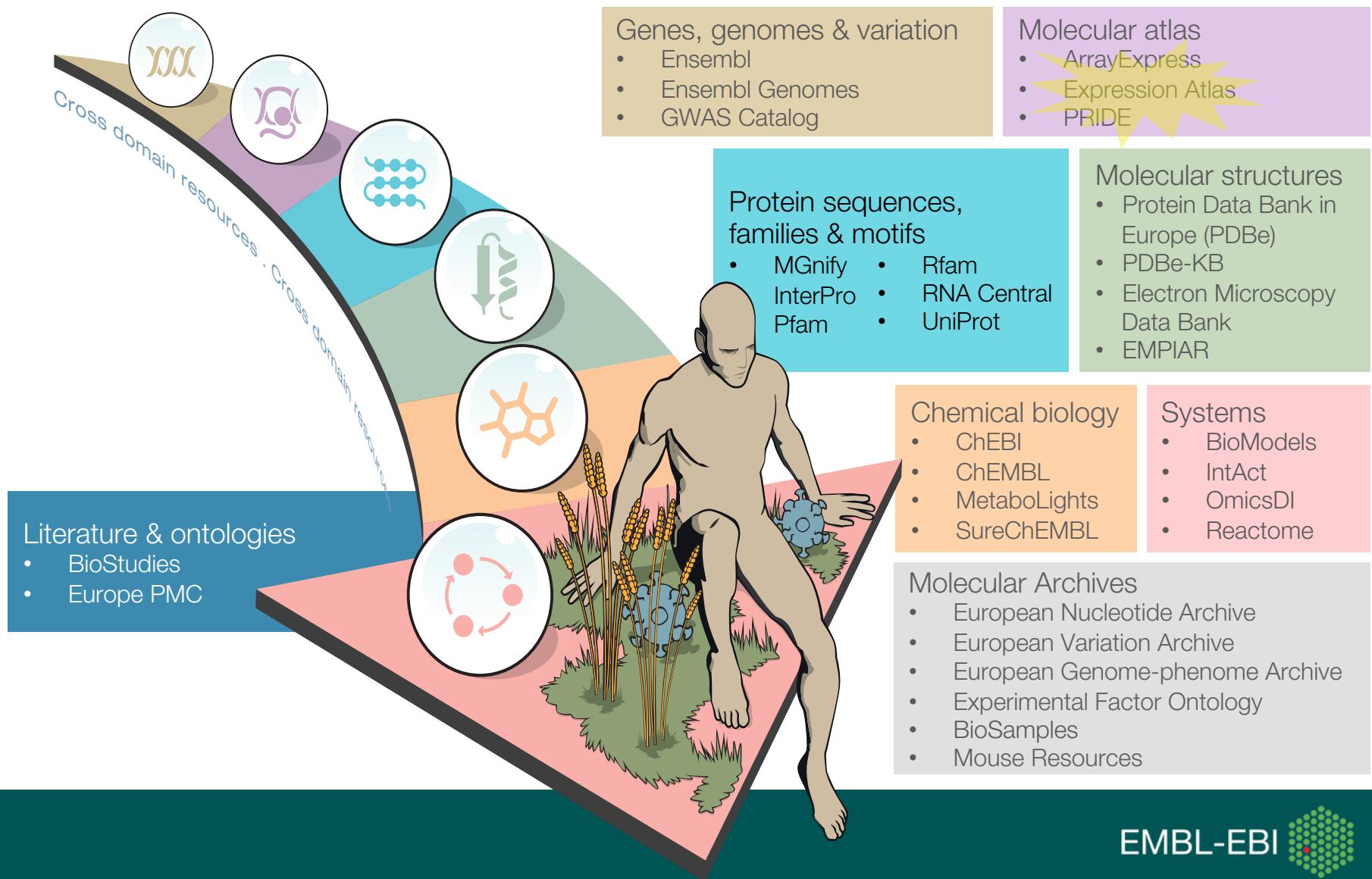
Share it with other data providers

Archive it

Classify it



Data resources at EMBL-EBI



Research

Leading computational biology
research since 1994

www.ebi.ac.uk/research

Research groups at EMBL-EBI



Alex Bateman Ewan Birney Pedro Beltrao Alvis Brazma Rob Finn Paul Flicek Andrew Leach Moritz Gerstung



Nick Goldman Zamin Iqbal John Marioni Evangelia Petsalaki Oliver Stegle Janet Thornton Virginie Uhlmann Daniel Zerbino

Postdoctoral fellowships at EMBL-EBI

Interdisciplinary

EIPOD projects bring together scientific fields that are usually separate, or transfer techniques to a novel context.

Genome Campus

ESPODs: joint Sanger–EMBL-EBI fellowships combine experimental and computational approaches.

Biomedicine

EBPODs are joint projects with EMBL-EBI, NIHR Cambridge Biomedical Research Centre and the University of Cambridge School of the Biological Sciences.

www.ebi.ac.uk/research/postdocs/

Training

Overview

Train at EMBL-EBI

Train outside EMBL-EBI

Train online

Webinars

About

10 Years

Contact

Bioinformatics training

EMBL-EBI Training

For scientists working at all levels

We train scientists at all levels to get the most out of publicly available biological data.

Follow us on Twitter and sign up to the EMBL courses and conferences newsletter.

www.ebi.ac.uk/training

Sponsoring 10 years of public training → Topic filtered by Gene expression [X]

Filter by: Topic

- biomodels
- DNA & RNA
- genomes
- Gene expression
- Proteins
- Structures
- Systems
- Ontologies
- Chemical biology
- Literature
- Cross domain

Filter by: Type

Train at EMBL-EBI

Train outside EMBL-EBI

Training events

May 2018

24

EMBL-EBI Workshop: Tools for functional genomics data

Conrad Hotel, Seoul, South Korea - 10 Gukjegeumyung-ro, Yeoeuido-dong, Yeongdeungpo-gu, Seoul, Seoul, South Korea

24th May

This course will provide participants with an introduction to functional genomics resources at EMBL-EBI, in particular to ArrayExpress and the submission tool Annotare.

Online courses

Start now

ArrayExpress: Discover functional genomics data quickly and easily

Author(s): Anja Füllgrabe

ArrayExpress is a database of functional genomics data. This course will give you an overview of how these data are stored in ArrayExpress and will teach you how to effectively search and retrieve data from the ArrayExpress website . Course...

Start now

ArrayExpress: Quick tour

Author(s): Melissa Burke



Celebrating 10 Years of Training

Building capacity: bioinformatics training

Train online

We provide free access to materials from EMBL-EBI courses, allowing you to choose when, where and how you learn

Webinars

Our collection of free seminars focuses on our range of data resources and new features and tools to boost your research

Courses at EMBL-EBI

Our on-site training courses help life-science researchers in all domains to make the most of public data resources

Courses at host sites

We send our expert trainers to host organisations throughout the world, where they provide training on bioinformatics

Industry engagement

Support and collaboration

www.ebi.ac.uk/industry

EMBL-EBI industry programme members

Pharmaceutical

abbvie



Celgene

Lilly

NOVARTIS

SANOFI

astellas
Leading Light for Life

Biogen

genzyme

MedImmune



Takeda

astex[®]
pharmaceuticals

**Boehringer
Ingelheim**

gsk

MERCK

Pfizer

ucb

AstraZeneca

Bristol-Myers Squibb

janssen
PHARMACEUTICAL COMPANIES
of Johnson & Johnson

MSD
Be well

Roche

Agri-food and consumer goods

Bayer CropScience

syngenta

Unilever



Elixir

Supporting the coordination of
biological data provision in Europe

www.elixir-europe.org



ELIXIR unites Europe's leading life science organisations in managing and safeguarding the data generated by publicly funded research.

It coordinates, integrates and sustains data resources across its member states, enabling scientists to access services that are vital for their research.



Navigating EBI – www.ebi.ac.uk

The home for big data in biology

Open data
EMBL-EBI shares data from life science experiments.

More about EMBL-EBI's impact in our annual report ➤
Data from 2016

Our unique Search service helps you explore dozens of biological data resources.
More about EBI Search ➤

All Find a gene, protein or chemical

Example searches: blast keratin bf1

Find a tool for your data analysis.

Find a tool ➤

Share your scientific data with the world.

Deposit data ➤

We are EMBL-EBI

The European Bioinformatics Institute (EMBL-EBI) is part of EMBL, Europe's flagship laboratory for the life sciences. More about EMBL-EBI and our impact ➤

Training

Access a wealth of world-leading training in bioinformatics and scientific service provision, regardless of your career stage or sector ➤

Data resources

Explore our open data resources to enrich your research. Browse data, perform analyses or share your own results. ➤

Industry

Explore our knowledge-exchange Industry Programme and take part in translational partnerships and projects ➤

Research

Find out about our research groups, postdoctoral schemes and PhD Programme ➤

ELIXIR

We support, as an ELIXIR node, the coordination of biological data provision throughout Europe ➤

Navigating EBI: tools & data resources

Services

Overview

A to Z

Data submission

Support

The European Bioinformatics Institute (EMBL-EBI) maintains the world's most comprehensive range of freely available and up-to-date molecular data resources.

Developed in collaboration with our colleagues worldwide, our services let you share data, perform complex queries and analyse the results in different ways. You can work locally by downloading our data and software, or use our [web services](#) to access our resources programmatically.

— You can read more about our services in the journal [Nucleic Acids Research](#)

Tools & Data Resources

Tools

Clustal Omega



Multiple sequence alignment of DNA or protein sequences. Clustal Omega replaces the older ClustalW alignment tools.

[Multiple sequence alignment](#)

Data resources

Ensembl

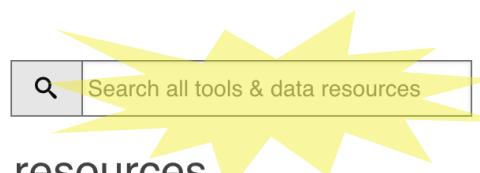


Genome browser, API and database, providing access to reference genome annotation

UniProt



A comprehensive resource for protein sequence and functional annotation.



Browse by type

XXX DNA & RNA	Gene Expression	Proteins
Structures	Systems	Chemical biology
Ontologies	Literature	Cross domain

Navigating EBI: tools & data resources

Overview

A to Z

Data submission

Support

Proteins | [View all tools and data resources](#) ›

Tools & Data Resources

Tools

BLAST [protein]



Fast local similarity search tool for protein sequence databases.

[Sequence similarity search](#)

ChEMBL BLAST Search

BLAST search against protein targets stored in ChEMBL database

[Sequence similarity search](#)

Clustal Omega



Multiple sequence alignment of DNA or protein sequences. Clustal Omega replaces the older ClustalW alignment tools.

Data resources

Complex Portal



A manually curated, encyclopaedic resource of macromolecular complexes from model organisms.

Enzyme Portal



Integrated enzyme data from EMBL-EBI resources.

IntAct



A curated database of molecular interactions.

IntEnz



Reference enzyme nomenclature and function.



Filter these tools & data resources

Browse by type

DNA & RNA	Gene Expression	Proteins
Structures	Systems	Chemical biology
Ontologies	Literature	Cross domain

Programmatic access

EMBL-EBI web services allow you to query our large biological data resources programmatically, so that you can develop data analysis pipelines or integrate public data with your own applications. The Web Services technology we use are built on open standards to ensure client and server software from various sources will work well together.

[Browse EMBL-EBI web services](#)

Navigating EBI: data submission

The home for big data in biology

Open data
EMBL-EBI shares data from life science experiments.

More about EMBL-EBI's impact in our annual report ➤
Data from 2016

Our unique Search service helps you explore dozens of biological data resources.
More about EBI Search ➤

All ▾ Find a gene, protein or chemical

Example searches: blast keratin bf1

Find a tool for your data analysis.
 Find a tool ➤

Share your scientific data with the world.
 Deposit data ➤

We are EMBL-EBI

The European Bioinformatics Institute (EMBL-EBI) is part of EMBL, Europe's flagship laboratory for the life sciences. More about EMBL-EBI and our impact. ➤

Training

Access a wealth of world-leading training in bioinformatics and scientific service provision, regardless of your career stage or sector ➤

Data resources

Explore our open data resources to enrich your research. Browse data, perform analyses or share your own results. ➤

Industry

Explore our knowledge-exchange Industry Programme and take part in translational partnerships and projects ➤

Research

Find out about our research groups, postdoctoral schemes and PhD Programme ➤

ELIXIR

We support, as an ELIXIR node, the coordination of biological data provision throughout Europe ➤

Navigating EBI: data submission

Services

Overview A to Z Data submission Support

SERVICES / DATA SUBMISSION

Data submission

Use this data submission wizard to find the right archive for your data in a few simple steps.

1 What **type of data** do you have?

DNA/RNA sequence

Expression data

Protein data

Structures

Systems

Chemical biology

Ontologies

Multi-omics or other cross-domain study

Why submit data to an archive?

Submission of primary data and derived information to public data repositories is an essential step in the scientific process. Through submission, the scientific community is fed the raw materials for the building and maintenance of the complete and up-to-date data sets that support searches and analysis on the latest sequences, structures and molecular profiles of living systems. Serving as a complement to the literature publication process and supporting early data sharing, the EBI offers a number of submission services appropriate for different types and scales of data.

All EMBL-EBI data repositories

[Array Express](#) ➤ functional genomics data

[BioModels](#) ➤ computational models

[BioSamples](#) ➤ reference sample data

[ChEBI](#) ➤ chemical entities

[DGVA](#) ➤ structural genetic variation data

[EFO](#) ➤ experimental variables

[EGA](#) ➤ human data that requires controlled access

[EMPIAR](#) ➤ raw image data

[ENA](#) ➤ nucleotide sequence data

[EVA](#) ➤ genetic variation data

[GO](#) ➤ Gene Ontology annotations

[IntAct](#) ➤ molecular interactions

[IntEnz](#) ➤ enzyme nomenclature

Navigating EBI: EBI search

The home for big data in biology

Open data
EMBL-EBI shares data from life science experiments.

More about EMBL-EBI's impact in our annual report ➤
Data from 2016

Our unique Search service helps you explore dozens of biological data resources.
More about EBI Search ➤

All ▾ LFY

Example searches: blast keratin bf1

Find a tool for your data analysis.

Share your scientific data with the world.

Deposit data ➤

We are EMBL-EBI

The European Bioinformatics Institute (EMBL-EBI) is part of EMBL, Europe's flagship laboratory for the life sciences. More about EMBL-EBI and our impact ➤

Training

Access a wealth of world-leading training in bioinformatics and scientific service provision, regardless of your career stage or sector ➤

Data resources

Explore our open data resources to enrich your research. Browse data, perform analyses or share your own results. ➤

Industry

Explore our knowledge-exchange Industry Programme and take part in translational partnerships and projects ➤

Research

Find out about our research groups, postdoctoral schemes and PhD Programme ➤

ELIXIR

We support, as an ELIXIR node, the coordination of biological data provision throughout Europe ➤

Navigating EBI: EBI search

EBI Search

Help & Documentation

About EBI Search



Feedback

Showing 27 results out of 14,012 in All results

Filter your results

Source

All results (14,012)

- Genomes & metagenomes (13)
- Nucleotide sequences (8,397)
- Protein sequences (3,266)
- Macromolecular structures (2)
- Gene expression (1,863)
- Molecular interactions (17)
- Protein families (7)
- Enzymes (1)
- Literature (325)
- Samples & ontologies (121)

LFY

Gene & protein summaries (includes expression, structures, literature...) (1 result)

-  [Floral meristem identity control protein LEAFY \(LFY\)](#)
-  [LFY \(LFY3, undefined\)](#)
-  [Thale Cress \(Arabidopsis thaliana\)](#)

Samples & ontologies (121 results)

[Kentrophoros sp. LFY](#) Source: Taxonomy (ID: 1905172)

Cross References: Nucleotide sequences (9) Samples & ontologies (1)

Formats: in ENA in Taxonomy format

[Candidatus Kentron sp. LFY](#) Source: Taxonomy (ID: 2126342)

Cross References: Nucleotide sequences (6) Samples & ontologies (1)

Formats: in ENA in Taxonomy format

[Neoplagioporus sp. LFY-1](#) Source: Taxonomy (ID: 2176032)

Cross References: Nucleotide sequences (2) Samples & ontologies (1)

Formats: in ENA in Taxonomy format

[View all 121 results for Samples & ontologies](#)

Navigating EBI: EBI search

Gene & protein summary for LFY

ORGANISMS
thale cress
Arabidopsis thaliana

Gene

Expression

Protein

Protein Structure

Literature

floral meristem identity control protein LEAFY (LFY)

[View in Ensembl Plants](#)

Gene Information and Sequence

- LFY spans 2742 bps of chromosome 5 from 24844248 to 24846989.
- LFY has 2 transcripts containing a total of 6 exons on the forward strand.
- Gene annotation by ARAPORT through a process of automatic and manual curation.
- [View the gene sequence in Ensembl Plants.](#)
- [View the chromosome region for this gene in Ensembl Plants.](#)

Variations

- LFY has 211 SNPs.
- [View sequence variations such as polymorphisms, along with genotypes and disease associations in Ensembl Plants.](#)

Orthologues

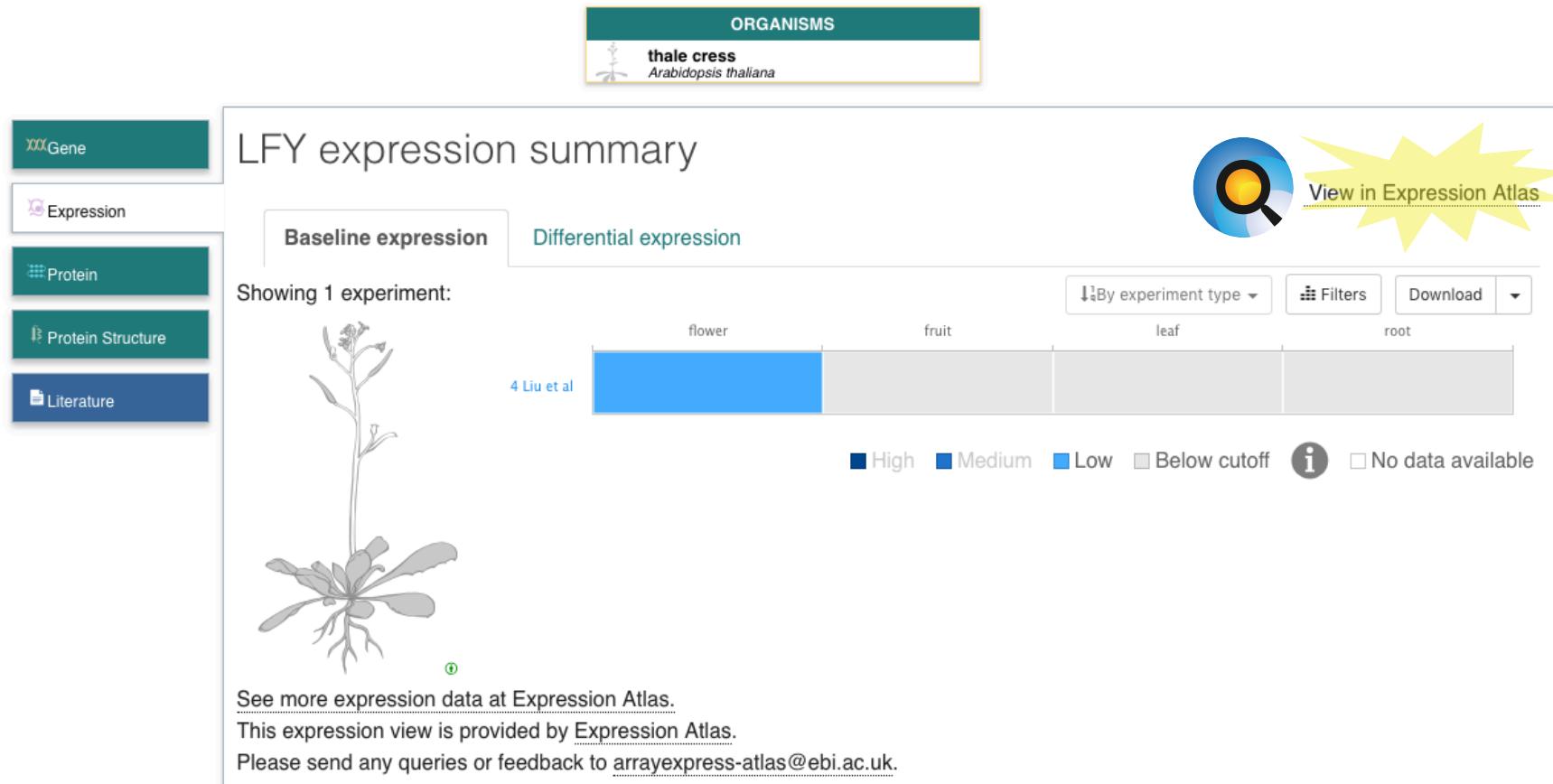
- LFY has 64 orthologues in Ensembl Plants
- [View homology between species inferred from a gene tree in Ensembl Plants.](#)

Paralogues

- LFY has no paralogues in Ensembl Plants

Navigating EBI: EBI search

Gene & protein summary for LFY



Navigating EBI: EBI search

Gene & protein summary for LFY

ORGANISMS

thale cress
Arabidopsis thaliana

Protein LEAFY

Functions

Probable transcription factor that promotes early floral meristem identity in synergy with APETALA1. Is required subsequently for the transition of an inflorescence meristem into a floral meristem, by an immediate upstream regulation of the ABC classes of floral homeotic genes. Activates directly APETALA1, CAULIFLOWER and AGAMOUS, and indirectly APETALA3 and PISTILLATA with the

[View annotation in UniProt](#)

Subcellular Location

Nucleus.

[View annotation in UniProt](#)

GO Cellular Component

nucleus

[View all in UniProt](#)

UniProt

View in UniProt

IPR035209

InterPro domain architecture

InterPro

Navigating EBI: EBI search

Gene & protein summary for LFY

ORGANISMS

thale cress
Arabidopsis thaliana

Gene

Expression

Protein

Protein Structure

Literature

2 protein structures available

Structure of LEAFY transcription factor from *Arabidopsis thaliana* in complex with DNA from AP1 promoter

Method

X-ray diffraction

Experiment

Resolution: 2.104Å

R-Factor: 22.59%

Free R-Factor: 25.15%

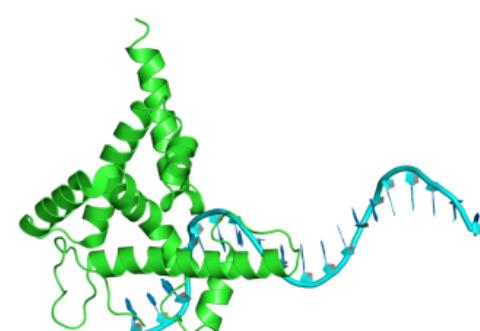
Dates

Deposited: 16-07-2008

Released: 23-09-2008

Revised: 07-08-2013

View in PDBe



Navigating EBI: EBI search

Gene & protein summary for LFY

ORGANISMS

thale cress
Arabidopsis thaliana

All Articles (1216)

LC-MS based sphingolipidomic study on A549 human lung adenocarcinoma cell line and its taxol-resistant strain.
(PMID:30089463 PMCID:PMC6083515)
Huang H, Tong TT, Yau LF, Chen CY, Mi JN, Wang JR, Jiang ZH.
BMC Cancer[2018, 18(1):799]
Cited:0 times

Clinical Value of Inflammation-Based Prognostic Scores to Predict the Resectability of Hyperbilirubinemia Patients with Potentially Resectable Hilar Cholangiocarcinoma.
(PMID:30076591)
Hu HJ, Jin YW, Zhou RX, Ma WJ, Yang Q, Wang JK, Liu F, Cheng NS, Li FY.
J Gastrointest Surg[2018]
Cited:0 times

Primary gastrointestinal stromal tumor of the liver.
(PMID:30077788)
Hu HJ, Fu YY, Li FY.
Clin Gastroenterol Hepatol[2018]

Reviews (122)

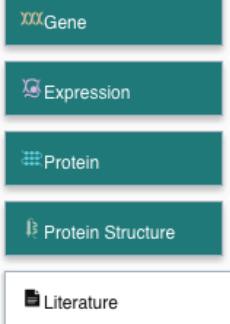
Revisiting the Role of Plant Transcription Factors in the Battle against Abiotic Stress.
(PMID:29857524 PMCID:PMC6032162)
Khan SA, Li MZ, Wang SM, Yin HJ.
Int J Mol Sci[2018, 19(6)]
Cited:0 times

Convergent Evolution and the Diverse Ontogenetic Origins of Tendrils in Angiosperms.
(PMID:29666627 PMCID:PMC5891604)
Sousa-Baena MS, Sinha NR, Hernandes-Lopes J, Lohmann LG.
Front Plant Sci[2018, 9:403]
Cited:1 times

The origin and early evolution of vascular plant shoots and leaves.
(PMID:29254961 PMCID:PMC5745332)
Harrison CJ, Morris JL.
Philos Trans R Soc Lond B Biol Sci[2018, 373(1739)]
Cited:0 times



Europe PMC



EBI is like a treasure trove of data



Janet Thornton, Director Emeritus, European Bioinformatics Institute



EMBL-EBI workshop: Exploring plant gene expression in Expression Atlas

Introduction to EMBL-EBI resources

Laura Huerta, PhD

Senior Scientific Curator

lauhuema@ebi.ac.uk

Copenhagen, 10 September 2018

