



PRECURSOR

CONSORTIUM
2024-2026

Coordination

Marie-Laure Martin
UMR IPS2

marie-laure.martin@inrae.fr

Sylvie Coursol

UMR IJPB

sylvie.coursol@inrae.fr

Key words

Plants

Cis-regulatory elements

Transposable elements

Knowledge base

Artificial intelligence

INRAE units involved

[IPS2](#)

[IJPB](#)

[URGI](#)

[MIA Paris-Saclay](#)

Partnerships

CIRAD

IRD

Université Clermont Auvergne

Expanding our fundamental knowledge of gene-proximal regions to improve selection models

Gene transcription is an essential process in the adaptive response of plants to environmental constraints. The interdisciplinary scientific consortium PRECURSOR aims to investigate and better understand how this process takes place in the proximal regions of genes to ultimately improve the predictive power of selection models.

Context and challenges

Transcription, the first stage of gene expression and protein synthesis, is tightly regulated by a number of molecular elements. *Cis*-regulatory elements, which consist of short DNA sequences, regulate gene expression via *trans*-acting factors that bind to the *cis*-regulatory elements.

Modifying gene expression through regulators

Cis-regulatory sequences are present in high density in the proximal regions of genes, but their characterization, an essential prerequisite for their use, remains incomplete. Recent projects have mapped DNA sequences preferentially located (known as PLMs) in these regions (in *Arabidopsis thaliana* and maize), with nearly 80% still unassigned in databases, although some are supported by MNase-defined cistrome occupancy analyses. Additionally,, numerous studies have shown that transposable elements (TEs) can include *cis*-regulatory sequences. When TEs are inserted near a gene, they can then affect the transcription of neighbouring genes by recruiting additional *trans* factors.

These two data sources (PLMs and TEs) are promising as they allow for the large-scale characterization of potential *cis*-regulatory elements. However, to gain a true understanding of proximal regions, these structural data need to be coupled with expression data. Original approaches using artificial intelligence may offer a promising way to integrate these biological data, thereby enabling the prediction of key genes and their regulatory networks.

However, there are few opportunities for teams of experts working in these areas to come together with their different and complementary skills. The PRECURSOR consortium was therefore established to overcome this obstacle by creating an interdisciplinary network of experts to address this topic.



© Freepik

Metaprogramme
DIGIT-BIO



digitbio@inrae.fr
http://www.inrae.fr/digitbio_eng/

Goals

PRECURSOR will bring together scientific teams working at the interface between biology (molecular science, genetics, physiology) and formal science (statistics, computer science, bioinformatics), to investigate different species (maize, wheat, sorghum) and gain a consolidated vision of the genetic basis for traits of agronomic interest that will encompass both structural and expression data.

The aim is to collaboratively advance the mapping and predictive power of *cis*-regulatory elements in the proximal regions of genes, taking into account the overall complexity of the question and the complementarities/differences between the species studied.

PRECURSOR's main objective is to form an interdisciplinary scientific consortium based on the unprecedented integration of heterogeneous data to gain a better understanding of the proximal regions of genes and ultimately to develop new alleles of agronomic interest and improve the predictive power of selection models.

Research units involved and partnerships

INRAE scientific division	INRAE research units	Expertises
<u>Plant Biology and Breeding</u>	<u>IPS2</u>	Bioinformatics of <i>cis</i> -regulatory elements, statistics of omics data
	<u>IJPB</u>	Biology of <i>cis</i> -regulatory elements; maize, environmental constraints, digestibility, functional genomics
	<u>URGI</u>	Information technology, knowledge bases, transposable elements
<u>Mathematics and Digital Technologies</u>	<u>MIA Paris-Saclay</u>	Artificial intelligence methods
External partners		Expertises
CIRAD	<u>AGAP</u>	Quantitative genetics, sorghum, functional genomics
IRD	<u>DIADE</u>	Biology, tropical cereals, root systems
University of Clermont Auvergne	<u>GDEC</u>	Molecular physiology of responses to biotic and abiotic stress, wheat, fungal pathogens, water stress

