Plant Science 252 (2016) 335–346

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Review article

# Seed maturation: Simplification of control networks in plants

Martine Devic*, Thomas Roscoe

*Régulations Epigénétiques et Développement de la Graine, ERL 3500 CNRS-IRD UMR DIADE, Centre IRD de Montpellier, 911 avenue Agropolis BP64501, 34394, Montpellier, France*

ARTICLE INFO

*Article history:*
Received 25 March 2016
Received in revised form 5 August 2016
Accepted 21 August 2016
Available online 22 August 2016

*Keywords:*
Seed maturation
AFL multigene family
Redundancy
Simplification
Minimal control network

ABSTRACT

Networks controlling developmental or metabolic processes in plants are often complex as a consequence of the duplication and specialisation of the regulatory genes as well as the numerous levels of transcriptional and post-transcriptional controls added during evolution. Networks serve to accommodate multicellular complexity and increase robustness to environmental changes. Mathematical simplification by regrouping genes or pathways in a limited number of hubs has facilitated the construction of models for complex traits. In a complementary approach, a biological simplification can be achieved by using genetic modification to understand the core and singular ancestral function of the network, which is likely to be more prevalent within the plant kingdom rather than specific to a species. With this viewpoint, we review examples of simplification successfully undertaken in yeast and other organisms. A strategy of progressive complementation of single, double and triple mutants of seed maturation confirmed the fundamental role of the AFL sub-family of B3 transcription factors as master regulators of seed maturation, illustrating that biological simplification of complex networks could be more widely applied in plants. Defining minimal control networks will facilitate evolutionary comparisons of regulatory processes and the identification of an essential gene set for synthetic biology.

## Contents

---

* Corresponding author.
*E-mail address:* martine.devic@ird.fr (M. Devic).

http://dx.doi.org/10.1016/j.plantsci.2016.08.012

## 1. Introduction

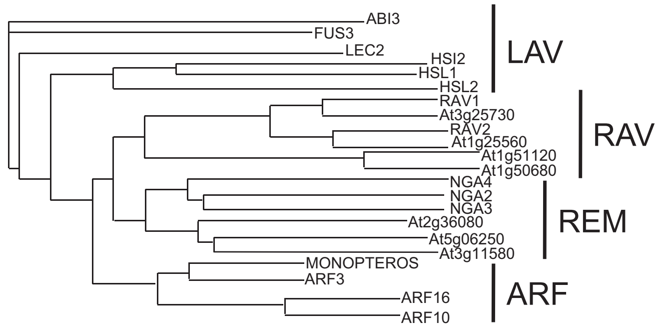
Seeds are an important component of human and animal nutrition supplying calories in the form of starch, sugar, and oil together with amino acids, vitamins and microelements. The synthesis of the seed reserves occurs after pattern formation of the embryo during maturation, the second phase of seed development. In the final phase of embryogenesis, desiccation tolerance is acquired and dormancy is established. An understanding of the control of the seed maturation phase is essential to improve seed quality traits. The biochemical pathways leading to the production of seed reserves in Arabidopsis have been genetically analysed. Several master regulators of seed maturation have been identified. Despite this insight, there is no simple model of the control of initiation and the progression through maturation. This is partly due to the redundancy among the master regulators and the multiple levels of regulation added during evolution. It is necessary to define the core components of the system rather than to describe its complexity in order to understand the gene regulatory network controlling seed maturation. A minimal control network may therefore represent the extant equivalent of an ancestral regulatory gene. Here, we highlight the necessity for simplification and the approaches to define minimal control networks.

## 2. Seed development and maturation

Seeds are complex structures, which have arisen in the Spermatophytes (seed bearing plants) more than 300 million years ago following whole genome duplication [1]. Gymnosperm seeds are composed of a diploid embryo (one maternal haploid set of chromosomes = 1 m and one paternal = 1 p) and a nourishing female gametophyte (2 m). Angiosperm seeds are the products of a double fertilisation and are composed of the seed coat, a diploid maternal tissue (2 m), the nutritive endosperm (often triploid 2m + 1 p) and the diploid embryo (1 m + 1 p), reviewed in [2]. The emergence of the seed trait has conferred advantages of protection of the reproductive structures and of dissemination as dry quiescent material. "Orthodox" Seeds are defined by their tolerance to desiccation (they can lose up to 90% of water during seed ripening). The embryo enters into a quiescent state and falls into dormancy concomitantly with the acquisition of desiccation tolerance. "Recalcitrant" seeds are less tolerant or are intolerant to desiccation, are less quiescent and cannot be well conserved, but they have the advantage of rapid germination. All seeds accumulate carbohydrate, lipid and protein reserves predominantly in the embryo of Gymnosperms and of exalbuminous Angiosperm seeds or in the endosperm of albuminous Angiosperm seeds. The formation of the embryo pattern, the accumulation of reserves, the acquisition of tolerance to desiccation and the entry into dormancy are the main programmes of seed development and are important agronomic traits defining seed quality. Collectively, these developmental processes comprise seed maturation. Irrespective of the qualitative differences listed above, all seeds essentially follow a common phase of seed maturation.

Genetic analyses in Arabidopsis have led to the identification of four master regulators of seed maturation: ABSCISSIC ACID INSENSITIVE3 (ABI3) [3], FUSCA3 (FUS3) [4], LEAFY COTYLEDON2 (LEC2) [5] and LEAFY COTYLEDON1 (LEC1) [6]. Whilst LEC1 is a subunit of the CCAAT binding complex (HAP3 or NF-YB9), a general eukaryotic transcriptional regulatory complex, ABI3, FUS3 and LEC2 (collectively named AFL) are transcription factors with a plant specific B3 DNA binding domain (Fig. 1). The AFL regulators influence most aspects of seed maturation. For example, the *abi3* and *fus3* mutants are intolerant to desiccation, *fus3* is not dormant (viviparous: can germinate on the mother plant) and *abi3*, *fus3* and *lec2* seeds contain less storage protein and lipid reserves and more starch.

**A. Phylogram of proteins with B3 domains**



**B. Phylogram of the B3 domain of LAV proteins**

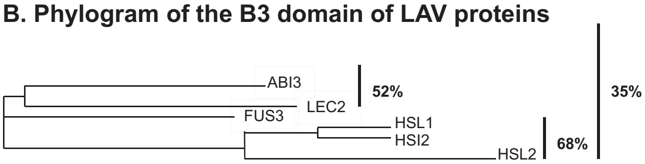


**Fig. 1.** Proteins possessing a B3 DNA binding domain in Arabidopsis.
A. Subset of B3 domain proteins. The Arabidopsis genome contains 118 genes encoding proteins possessing a B3 DNA binding domain. The amino acid sequences of a selection of 22 proteins containing a B3 DNA binding domain were used to construct a phylogram using ClustalW and ClustalTree (ebi.ac.uk). These proteins can be divided into four major families. The LAV (LEC-ABI3-VAL) family contains the three AFL and the three HSI/VAL with the additional EAR (ERF-associated amphiphilic repression) domain. The RAV (Related to ABI3/VP1) family has 13 members, first identified by their homology to VP1. Certain RAV proteins possess an additional AP2 domain. The ARF (Auxin Response Factor) family has 23 members. ARF1, the founder member, was identified as binding upstream several auxin response genes. The REM family (REproductive Meristem) has 76 members that can be subdivided into 6 subgroups. Most of the REM proteins contain more than one B3 DNA binding domain. The graphic illustrates that the nearest neighbour proteins to the AFL are the HSI/VAL proteins. Members of the three other families, ARF, RAV and REM, constitute separate clades.
B. Relatedness of the amino acid sequences of the B3 domain of AFL and HSI/VAL transcription factors. The numbers indicate the percentage of amino acid identity of the DNA binding domain only.

Homologues of AFL proteins have been identified in the genomes of all seed plants sequenced to date and in some moss and algae. It is hypothesised that their function in seed maturation is conserved among Spermatophytes [7–9]. Defining the core roles of AFL in the model plant Arabidopsis will provide insight as to regulation of the maturation phase of embryogenesis in cultivated crops.

## 3. AFL proteins in the plant kingdom

Among the 118 proteins in Arabidopsis that contain a B3 DNA binding domain, ABI3, FUS3 and LEC2 have the most conserved B3 domain and are more similar to each other than to any other B3 protein [10] (Fig. 1). Analysis of individual AFL loss of function phenotypes has demonstrated the existence of a strong redundancy in their mode of action together with a limited specificity, that is, specialisation of function [11].