Kerstin Kaufmann, Cezary Smaczniak and Gerco Angenent

The transcription machinery underlying flower formation

A key step in transcription is the sequence specific binding of transcription factors to their DNA recognition sites. In plants virtually no information is available about the composition of transcription factor complexes. Also DNA recognition and binding is still a mystery. Here we present the first results of the composition of transcription factor complexes and protein-DNA interactions by *in vivo* measurements. A master role is reserved for the glue protein SEP3, which is required for the formation of the flower and all its organs.

A key developmental process in multicellular organisms is the specification of identity of cells, which is the first step of differentiation into a distinct tissue and organ. Prominent examples of identity-specification proteins are the Hox transcription factors in animals that direct the anterior-posterior patterning. In plants, members of the MADS domain family of transcription factors facilitate this identity function and play a crucial role in the specification of floral organ identity. These floral organ identity factors have been well studied, particularly at the genetic level. Four classes of MADS domain proteins, A, B, C and E act together to form the four types of floral organs, sepals, petals, stamens and carpels, while a fifth class (class D) determines, in combination with the C and E classes, the identity of the ovule in which the female reproductive cells are formed (see Figure 1A). These MADS box domain proteins are supposed to dimerise into homo- or heterodimers that bind to a DNA consensus called the CArG box (consensus motif is $CC(A/T)_6GG$). Based on genetic information and yeast-based interaction studies, it has been proposed that these MADS proteins form tetrameric complexes composed of a double dimer. According to the so-called 'floral quartet-model [1], these tetrameric complexes bind to two adjacent recognition sites (CArG boxes), which results in DNA bending to facilitate transcription regulation (see Figure 1B). Main research questions in our group are (i) what is the composition of these floral organ identity complexes and (ii) how do they recognise specific DNA sequences and control the expression of down-stream target genes?

The glue protein SEP3 To investigate the composition of the MADS domain dimers and possible tetrameric complexes, we performed comprehensive yeast two- and three-hybrid studies using all available Arabidopsis MADS proteins. The Arabidopsis genome encodes 107 members of this transcription factor family which are not only involved in the control of floral organ identity, but also in many other developmental processes, such as root development, flower initiation, and fruit formation. The yeast two hybrid study [2] yielded more than 250 different possible dimer combinations. Many of these combina-

What this research is about:

From protein to flower

Nature has a large biodiversity of flowers, and not without purpose. For plants this is of vital importance since the flower contains the reproductive organs. During evolution flowers have adapted so that they are the most attractive for insects and best equipped for fertilization. "All the flower's organs are essential for reproduction. The plant life cycle depends on it. This knowledge is highly interesting for seed improvement and agriculture applications too," explains Gerco Angenent of the Business Unit Bioscience at Plant Research International. "The formation of flowers is regulated at the genetic level by transcription factors and the genes they regulate. But still little is known about how these transcription factors act together and how they function in the plant."

Although the variation in flower shapes and colours are numerous, four basic organs are generally the same in every flower: a flower contains sepals to protect the floral bud, colorful petals to seduce insects and the two reproductive organs. Angenent: "By studying spontaneous mutants of the *Arabidopsis* plant, for instance plants which produce flowers with only one organ type, we can learn which genes are switched on and off and which transcription factors are involved. *Arabidopsis* is a frequently used and fast growing member of the cabbage family similar to the cauliflower."

In this article the authors describe the unraveling of a network of transcription factors leading to the formation of the flower organs. They also discovered a master regulator among the transcription factors: SEP3. This key protein is involved in many transcription factor protein complexes and therefore directly or indirectly responsible for gene regulation in the flower. "This transcription factor is at the top of the pyramid and initiates a cascade of processes that finally forms the flower organs. On top of that, it is part of many protein complexes in the pyramid too," says Angenent. "How the other accompanying transcription proteins are involved in gene regulation and development of flowers is the next question we have to solve."

NPC Research Theme T2: Proteome Biology of Plants

tions were confirmed by genetic studies or other biochemical approaches (e.g. by co-immunoprecipitation or Fluorescence Resonance Energy Transfer-FRET-based approaches). Based on the yeast data we were able to build an algorithm for the identification of small peptide motifs (4-10 AA in length) in the two dimerising proteins that determine the specificity of protein interaction [3]. Using these motifs and the 'predictor' algorithm, we were able to predict which MADS domain proteins are interacting with each other.
In the following experiment, we performed a yeast 3-hybrid experiment starting with the dimers obtained from the 2-hybrid experiment and the about 100 single MADS proteins.

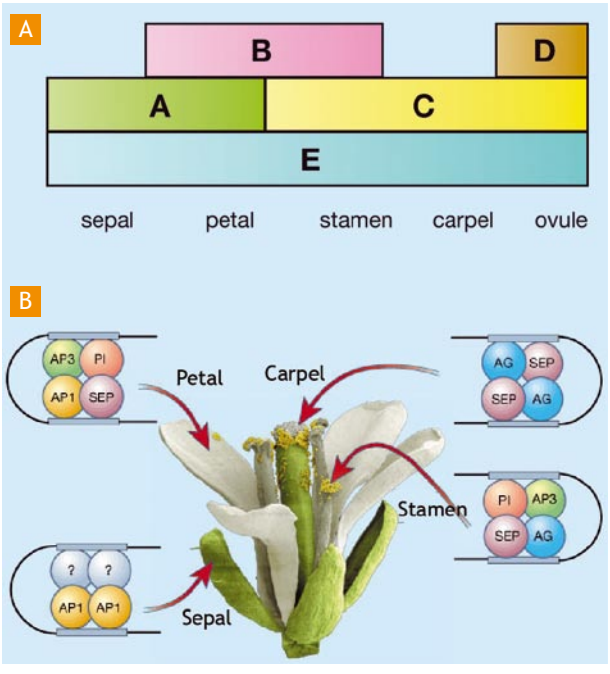


Figure 1 | MADS domain proteins in floral organ development
A: ABC model describing the classes of genes that are responsible for identity specification of the floral organs, sepals, petals, stamens, carpels and ovules. The A function leads to sepal formation; A+B give rise to petals; B+C, stamens; C alone controls carpel identity and the D function is required for ovule formation. The E function gene SEP3 is involved in the formation of all organs.
B: Quartet model: 4 MADS domain proteins form a tetrameric complex that binds to two binding motifs in the DNA.

The limitation of this experiment is that some proteins contain an activation domain, which activates the yeast reporter gene without interacting with the other protein ('auto-activation'). Nevertheless, we obtained over 100 distinct ternary complex combinations of which the majority contains the class E protein SEPALLATA3 (SEP3). This protein is therefore designated as the 'glue protein' and appeared to be present in most of the complexes required for the formation of the flower and its organs [4].

The next question we addressed ourselves was whether the MADS domain transcription factors are able to interact with other types of transcription factors. To get an initial idea about the possible existence of protein complexes with a mixed composition of several transcription factor classes, we performed a yeast 3-hybrid using a MADS dimer as bait (a class C-E combination) and the almost complete Arabidopsis transcription factor collection (1200 TFs) as prey. This experiment revealed for the first time interactions between MADS transcription factors and transcription factors of the bHLH and plant-specific TCP families. What these interactions mean in terms of biological consequences await further genetic and biochemical studies.

In vivo dynamics To uncover the composition of the MADS domain protein complexes that are active in the flower, we performed immunoprecipitation experiments using antibodies directed against SEP3. Floral tissues were harvested and the immunoprecipitated sample was digested with trypsine and subsequently, the peptides were analysed by LC-MS/MS (LTQ-MS/MS) in collaboration with Martina O'Flaherty from the group of Prof. Albert Heck. These analyses revealed that SEP3 interacts *in vivo* with MADS domain members of the A, B, C and D classes (see Figure 1), all involved in the formation of floral organ formation. In fact, all MADS domain proteins that were previously identified by yeast 2/3-hybrid could be confirmed by these MS-studies. Whether these proteins are components of quaternary complexes as the quartet model [1] proposes or that they form separate dimer pairs containing SEP3 need further analysis.

Because SEP3 is involved in the formation of all organs and we would like to unravel the complexes involved in the separate organs, we will use other baits (e.g. AP1 for the sepals and petals) and make use of specific homeotic mutant plants that produce flowers with only one organ type (see Figure 2A). We also expect that these transcription factor complexes are dynamic in their composition and therefore, we aim in our NPC2 project to analyse the complexes during the development of the floral organs from onset onwards. For this, a specific mutant (see Figure 2B) will be used in which no organs are formed until an inducible form of AP1 is activated [5] by exogenously applying a steroid hormone. Induction of AP1 leads to the synchronized initiation of floral organ in the cauliflower-like inflorescence.

Auto regulatory network It is known that the MADS domain proteins control the identity of floral organs by activating or suppressing the transcription of target genes. SEP3 is a major regulator of transcription in the flower, although limited information is available, which genes are controlled and how SEP3 binds to its target DNA. To identify the target genes with a genome-wide approach, we performed chromatin immunoprecipitation (ChIP) experiments. In such an approach, the transcription factor that is bound to the DNA is precipitated. After precipitation, the DNA that was co-precipitated with the transcription factor is

Figure 2 | Two floral mutants used in our research.
A. The agamous (ag) mutant with reproductive organs replaced by petals.
B. The 'cauliflower' mutant with numerous branched floral meristems, all in approximately the same developmental stage.

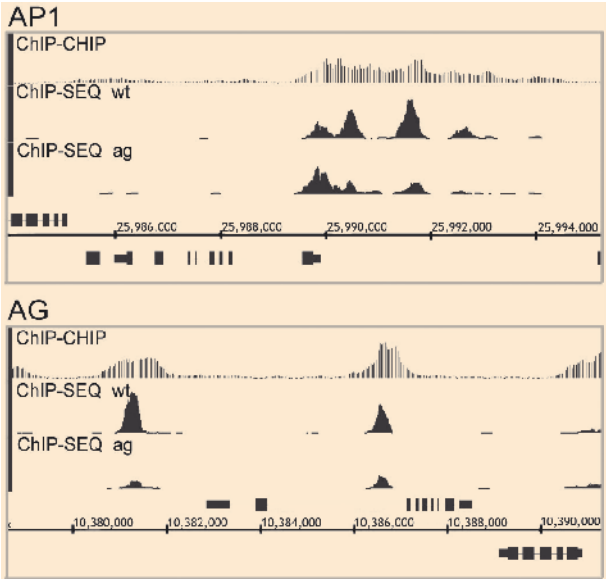


Figure 3 | Examples of SEP3 binding sites in other MADS box genes (AP1 and AG). The peaks are tiled sequences derived from the ChIP-SEQ experiment and represent binding sites. The gene structure is represented at the bottom of the graph.

deep-sequenced by the Illumina (Solexa) platform. Peaks of piled sequences represent binding sites in the genome (see Figure 3). This ChIP-SEQ technology was used to identify target genes of SEP3 (E function)[6] and AP1 (A function).

It appeared that these key regulators bind to thousands of sites in the Arabidopsis genome and, as expected, the fast majority of the binding sites contains the consensus binding site for MADS domain factors (CArG box). Among the targets are many MADS box genes, which points to the existence of an autoregulatory network, where MADS domain factors activate or suppress the transcription of other members of the family. Also other regulators, such as transcription factors and microRNAs are among the targets.

Surprisingly we also observed that many genes involved in the auxin signalling pathway are under the control of SEP3. Auxin is a plant hormone that plays a key role in patterning of organ primordia, growth and differentiation of plant organs. This novel observation links the MADS transcriptional machinery in the flower with the action of a key regulator of developmental processes, the growth regulator auxin.

Also with respect to the identification of target genes and the initiation of organ formation, we aim for an approach that enables a dynamic view of the regulatory pathway. To achieve this, we will use the cauliflower mutant (see Figure 2B) again and induce floral organ formation by applying an AP1 inducing agent.

References

1 Theissen, G. and Saedler, H. (2001) Floral quartets. *Nature* 409: 469-471

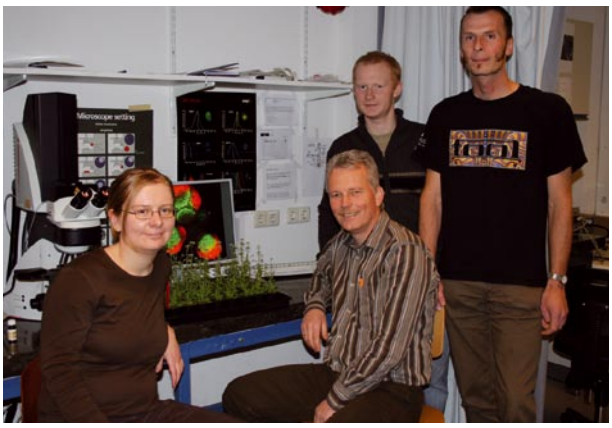
2 de Folter, S. *et al*. (2005) Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. *Plant Cell* 17: 1424-1433.

3 van Dijk, A.D.J. (2008). Predicting and understanding transcription factor interactions based on sequence level determinants of combinatorial control. *Bioinformatics* 24, 26-33.

4 Immink, R.G.H *et al*. (2009) SEPALLATA3: The "glue" for MADS box transcription factor complex formation. *Genome Biology*, 10(2): R24.

5 Wellmer, F. *et al*. (2006) Genome-Wide Analysis of Gene Expression during Early Arabidopsis Flower Development. *PLoS Genet* 2(7): 1012-1024.

6 Kaufmann, K. *et al*. (2009) Target genes of the MADS transcription factor SEPALLATA3: Integration of developmental and hormonal pathways in the Arabidopsis flower. *PLoS Biology*, in press.



Research team

The research team working on MADS domain interactions and complexes. From left to right: Kerstin Kaufmann, Gerco Angenent, Cezary Smaczniak and Richard Immink.

Contact

Prof. Gerco Angenent
Plant Research International
Bornsesteeg 65
6708 PD Wageningen
The Netherlands
T +31 317 480 953
Gerco.angenent@wur.nl

Summary

MADS domain transcription factors play a key role in the specification of floral organ identity. They act in a combinatorial way with other MADS proteins and transcription factors of other families. Presumably they bind to DNA of the target genes as tetrameric MADS complexes. The composition of the transcription factor complexes determines to a large extent the DNA binding specificity and the activity (i.e. transcriptional activator or suppressor) of the transcription complex. Using yeast-based assays we have determined the protein-protein interaction network among the MADS family members. It appeared that SEP3 is present in many of these transcription factor complexes and as such, it is a key regulator that regulates molecular pathways leading to the initiation and further differentiation of all organs in the flower. Using immunoprecipitation followed by MS analysis of the protein complex components, it was demonstrated that SEP3 interacts in vivo with all other MADS domain organ identity factors in the flower. To unravel the transcriptional pathway controlled by SEP3, we used a chromatin immunoprecipitation approach followed by sequencing of the bound DNA (ChIP-SEQ). This yielded a genome-wide collection of direct target genes and shed light on how SEP3 control identity, growth, and hormonal action in the flower.

summary