TECHNICAL ADVANCE

Genome-scale Arabidopsis promoter array identifies targets of the histone acetyltransferase GCN5

Moussa Benhamed[1], Marie-Laure Martin-Magniette[2,3], Ludivine Taconnat[2], Frédérique Bitton[2], Caroline Servet[1], Rebecca De Clercq[4,5], Björn De Meyer[4,5], Caroline Buysschaert[4,5], Stéphane Rombauts[4,5], Raimundo Villarroel[4,5], Sébastien Aubourg[2], Jim Beynon[6], Rishikesh P. Bhalerao[7], George Coupland[8], Wilhelm Gruissem[9], Frank L.H. Menke[10], Bernd Weisshaar[11], Jean-Pierre Renou[2], Dao-Xiu Zhou[1,*] and Pierre Hilson[4,5,*]

[1]*Institut de Biotechnologie des Plantes, UMR 8618, Centre National de la Recherche Scientifique, Université de Paris Sud 11, 91405 Orsay, France,*
[2]*Unité de Recherche en Génomique Végétale, UMR INRA 1165-CNRS 8114-UEVE, 91057 Evry Cedex, France,*
[3]*Unité de Mathématiques et Informatique Appliquées, UMR AgroParisTech-INRA, 75231 Paris Cedex, France,*
[4]*Department of Plant Systems Biology, Flanders Institute for Biotechnology, 9052 Gent, Belgium,*
[5]*Department of Molecular Genetics, Ghent University, 9052 Gent, Belgium,*
[6]*Warwick Horticultural Research International, University of Warwick, Warwick CV35 9EF, UK,*
[7]*Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, S-901 83 Umeå, Sweden,*
[8]*Abteilung Entwicklungsbiologie der Pflanzen, Max-Planck-Institut für Züchtungsforschung, 50829 Köln, Germany,*
[9]*Institute of Plant Sciences, Eidgenössische Technische Hochschule Zurich, 8092 Zurich, Switzerland,*
[10]*Molecular Genetics Group, Department of Biology, Utrecht University, 3584 CH Utrecht, The Netherlands, and*
[11]*Department of Biology, Genome Research, Bielefeld University, 33594 Bielefeld, Germany*

Summary

We have assembled approximately 20 000 *Arabidopsis thaliana* promoter regions, compatible with functional studies that require cloning and with microarray applications. The promoter fragments can be captured as modular entry clones (MultiSite Gateway format) via site-specific recombinational cloning, and transferred into vectors of choice to investigate transcriptional networks. The fragments can also be amplified by PCR and printed on glass arrays. In combination with immunoprecipitation of protein–DNA complexes (ChIP–chip), these arrays enable characterization of binding sites for chromatin-associated proteins or the extent of chromatin modifications at genome scale. The Arabidopsis histone acetyltransferase GCN5 associated with 40% of the tested promoters. At most sites, binding did not depend on the integrity of the GCN5 bromodomain. However, the presence of the bromodomain was necessary for binding to 11% of the promoter regions, and correlated with acetylation of lysine 14 of histone H3 in these promoters. Combined analysis of ChIP–chip and transcriptomic data indicated that binding of GCN5 does not strictly correlate with gene activation. GCN5 has previously been shown to be required for light-regulated gene expression and growth, and we found that GCN5 targets were enriched in early light-responsive genes. Thus, in addition to its transcriptional activation function, GCN5 may play an important role in priming activation of inducible genes under non-induced conditions.

Keywords: Arabidopsis, chromatin immunoprecipitation, histone acetylation, ChIP–chip, promoter, Gateway site-specific recombination.

Introduction

By affecting nucleosome positioning and density, histone modifications play a pivotal role in chromatin remodeling and gene regulation (Horn and Peterson, 2002). Covalent additions to the N-terminal tails of the core histones include

acetylation, methylation, phosphorylation, ADP ribosylation and ubiquitinylation. These modifications dynamically shape chromatin structure in various, sometimes opposite, ways. Their pattern determines the open or repressed state of the cognate DNA sequences (Jenuwein and Allis, 2001). In particular, hyper-acetylation relaxes chromatin structure and is associated with transcriptional activation, whereas hypo-acetylation induces chromatin compaction and gene repression (Millar and Grunstein, 2006). Histone acetylation also provides epigenetic markers for gene expression because it blocks association of heterochromatin-stabilizing complexes.

The homeostatic balance of nucleosomal histone acetylation is maintained by the antagonistic action of histone acetyltransferases (HATs) and histone deacetylases (HDACs). HAT and HDAC proteins usually function in multi-protein complexes. For instance, the yeast HAT protein Gcn5 is the catalytic subunit of the Spt-Ada-Gcn5 acetyltransferase (SAGA), SAGA-like (SLIK), transcriptional adaptor (ADA) and HAT-A2 complexes that have distinct histone targets and effects on transcriptional activity (reviewed in Lee and Workman, 2007). In addition to its role in histone acetylation, Gcn5 primarily functions as a key subunit of transcriptional co-activator complexes.

Chromatin regulators, including HAT complexes, have no sequence-specific DNA recognition properties of their own and are recruited to specific locations in the genome by various mechanisms (Lee and Workman, 2007). First, GCN5 association with promoters can be mediated directly or indirectly by DNA-binding transcription factors (Robert et al., 2004). For example, in yeast, Gcn5 association with the *HO* promoter depends on the transcription factor SBF (Cosma et al., 1999). In other cases, GCN5 is bound to promoters through interaction between GCN5-containing complexes, such as ADA or SAGA, and promoter-binding transcription factors. In *Arabidopsis thaliana*, the cold-responsive transcription factor CBF1 was recently shown to interact with ADA2, which in turn binds GCN5 (Mao et al., 2006). Second, the GCN5-containing SAGA complex can be recruited to promoters by recognizing specific histone markers, such as the trimethylated H3K14 residue (Pray-Grant et al., 2005). Third, the GCN5 bromodomain binds to acetylated histone tails, thereby mediating GCN5 association with chromatin for propagation of general histone acetylation (Imoberdorf et al., 2006; Mujtaba et al., 2007).

The Arabidopsis HAT family contains 12 members, which fall into four groups based on primary homology with yeast and mammalian HATs: GCN5-related N-acetyltransferase (GNAT), MOZ-YBF2/SAS3-SAS2-TIP60 (MYST), CREB-Binding Protein (CBP) and TBP-associated factor 1 (TAF1) (Pandey et al., 2002). Among the GNAT members, GCN5 (or HAG1) has been functionally characterized. GCN5 is evolutionarily conserved, with orthologs in other eukaryotes, and contains a HAT N-terminal catalytic domain and a C-terminal bromodomain (Pandey et al., 2002). Deletion mutations of *GCN5* result in either up or downregulation of 4% of the genes in yeast and cause embryo growth defects in mice (Lee et al., 2000; Yamauchi et al., 2000). In *Drosophila melanogaster*, null *dGcn5* alleles block the onset of both oogenesis and metamorphosis (Carré et al., 2005). T-DNA insertion mutations of Arabidopsis *GCN5* induce pleiotropic defects that alter plant development and responses to environmental conditions, such as light and cold (Benhamed et al., 2006; Bertrand et al., 2003; Vlachonasios et al., 2003), suggesting that GCN5 is involved in both the long-term regulation of epigenetic chromatin landscapes and the control of short-term transcriptional switches. Mutations in *GCN5* affect the expression of approximately 4% of the Arabidopsis genes in seedlings (Vlachonasios et al., 2003). GCN5 is involved in the acetylation of both histone H3 and H4 near the proximal promoter region of several light-inducible genes, suggesting that it is recruited by DNA-binding transcription factors (Benhamed et al., 2006).

Although GCN5 plays an essential role in chromatin modification and transcriptional regulation, its mode of action is still not understood. Therefore, it is important to determine which sites are targeted by GCN5 in the Arabidopsis genome in order to investigate the mechanisms it regulates. In recent years, whole-genome array technologies have been used to identify the sites bound by chromatin-associated proteins *in vivo* (Hanlon and Lieb, 2004) or targeted by chromatin-modifying enzymes (Millar and Grunstein, 2006; Zhang et al., 2006). We used chromatin immunoprecipitation (ChIP) in combination with a novel Arabidopsis promoter array to demonstrate that GCN5 associates with a large number of transcriptionally active promoters. GCN5–promoter binding appeared to be mostly independent of the bromodomain that recognizes acetylated histone tails (reviewed in Mujtaba et al., 2007). We found, however, that the bromodomain is required for binding and acetylating lysine 14 of histone H3 on a subset of promoter regions, 10% of which are promoters of transcription factor genes. Considering light-regulated genes, GCN5 target promoters are enriched in early light-responsive genes.

Results

The Arabidopsis SAP promoter repertoire

To facilitate the study of transcriptional networks in Arabidopsis, we developed a reference repertoire of Arabidopsis promoter sequences compatible with both microarray applications and downstream functional studies that require cloning (Figure 1). The systematic analysis of Arabidopsis promoters (SAP) repertoire was constructed
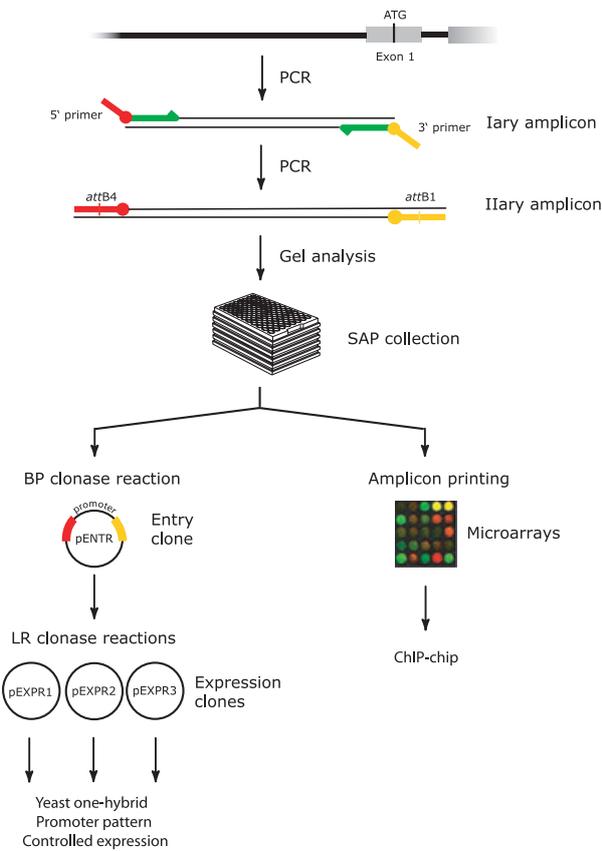
Figure 1. Construction and exploitation of the SAP promoter repertoire.

in two successive rounds. The first phase focused on gene models supported by experimental full-length cDNA sequences available in GenBank (December 2003), and included an *in silico* set of 11 893 promoter amplicons (SAPv1). The second phase extended the repertoire to all gene models whose transcription could be detected in the AtGenExpress ATH1 genechip profile compendium (http://www.arabidopsis.org/info/expression/ATGenExpress. jsp), complemented by data obtained by Yamada *et al.* (2003), resulting in the addition of 10 367 promoter amplicons, giving a total set of 22 260 *in silico* amplicons (SAPv2).

Each promoter amplicon was defined *in silico* by a pair of oligonucleotides designed for PCR synthesis from genomic DNA of the Arabidopsis Columbia-0 ecotype. The amplicon 3′ end was the nucleotide preceding the translation initiation codon in the corresponding gene model. Because multiple transcription initiation sites can be defined for individual genes, the amplified fragments also included the corresponding 5′ untranslated region, although here they are referred to as promoters. The amplicon 5′ end was located at the boundary between the targeted intergenic region and the adjacent upstream open reading frame

(ORF), or approximately 2.0 kb upstream of the translation initiation codon.

SAP database

Promoter amplicon information is available via the SAP online database at http://www.psb.ugent.be/SAP/. For each amplicon, the database provides its sequence (as deduced from the Arabidopsis genome), PCR primers, length, genome location, percentage GC content, corresponding gene code, and additional technical information. The relevant information can also be downloaded as a tab-delimited dump file (http://www.psb.ugent.be/SAP/). The SAP promoter repertoire has been integrated into the FLAGdb++ genome viewer (http://urgv.evry.inra.fr/FLAGdb; Samson *et al.*, 2004), in which details about individual sequences can be accessed by clicking on their graphical representation.

Promoter amplicon synthesis

In total, 19 741 promoter amplicons were successfully synthesized, as verified by the detection of a single PCR product of the expected size by agarose gel electrophoresis. Most other PCRs did not yield a product that was detectable in the gel because of the amplicon size. Sequences of 2.0 kb or smaller are preferred, because larger fragments yielded less robust PCR amplifications. All primary amplification results are documented in the SAP database, including links to agarose gel images and 96-well microtiter plate maps with PCR results. Promoter sequences were amplified in a two-step PCR for addition of universal extensions carrying the *att*B1 and *att*B4 Gateway cloning sites (Figure 1).

During the first phase of the SAP project, the SAPv1 collection (consisting of 124 96-well plates) was re-amplified in a subsequent PCR. Agarose gel analysis of PCR products confirmed the synthesis of 8471 probes of the expected size and in sufficient amount. Below, we present two SAPv1 array applications for studying the association between chromatin and remodeling enzymes and the acetylation status of core histones.

Genome-wide association of GCN5 with target promoters

Arabidopsis promoters bound by GCN5 *in vivo* were identified in ChIP–chip assays. A purified polyclonal antibody used to immunoprecipitate chromatin fragments was raised against two GCN5 peptides located in the N-terminal portion of the protein (Benhamed *et al.*, 2006). The antibody detected a protein of 62 kDa in the wild-type, as expected for the GCN5 protein, but smaller polypeptides in the *gcn5-1* and *gcn5-2* mutants that each carry a T-DNA insertion in the 3′ end of the gene (Figure S1). The T-DNA insertion in *gcn5-1*
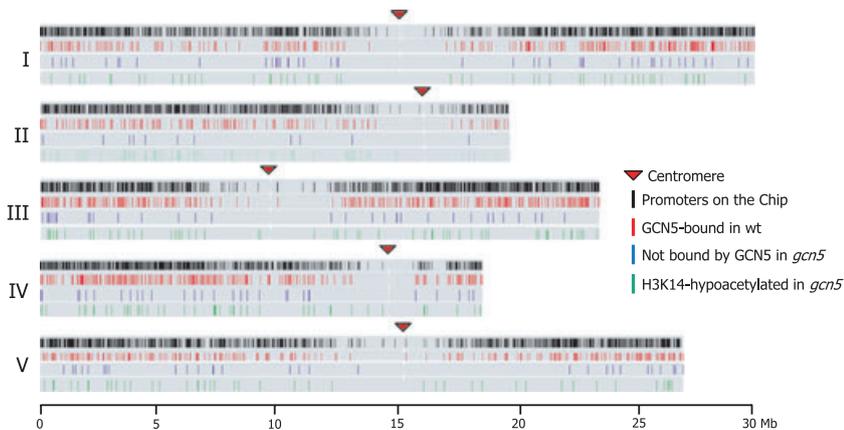
Figure 2. Genome-wide distribution of Arabidopsis GCN5 targets. From top to bottom and for each Arabidopsis promoter (I–V): red arrowheads, centromeres; black, position of promoter sequences spotted on the SAPv1 array; red, GCN5 binding promoter targets; blue, promoters for which GCN5 binding was abolished in the *gcn5-1* and *gcn5-2* mutants; green, promoters that were H3K14 hypo-acetylated in the *gcn5-1* and *gcn5-2* mutants.

▼ Centromere
▮ Promoters on the Chip
▮ GCN5-bound in wt
▮ Not bound by GCN5 in *gcn5*
▮ H3K14-hypoacetylated in *gcn5*

removes 17 codons of the C-terminal bromodomain (Vlachonasios *et al.*, 2003), whereas the insertion in *gcn5-2* results in complete deletion of the bromodomain (Bertrand *et al.*, 2003; Long *et al.*, 2006).

GCN5-associated chromatin fractions were extracted from wild-type, *gcn5-1* and *gcn5-2* plants, and three biological ChIP–chip replicates were analyzed for each of them. The DNA purified in each ChIP (defined as the IP fraction) was amplified, labeled, and hybridized to the promoter array. After data normalization to remove technical biases, the IP hybridization signal histogram was modeled using a mixture model: components of the mixture were interpreted to identify 'enriched' probes corresponding to GCN5-associated promoter regions (see Experimental procedures).

ChIP–chip results showed that, in wild-type plants, GCN5 binds to a large proportion of the promoters contained in the SAPv1 promoter repertoire (3352/8471, i.e. 40%; Table S1). Target promoters were distributed over all five chromosomes (Figure 2). Chromatin isolated from *gcn5-1* and *gcn5-2* plants had comparable profiles (Pearson correlation = 0.78, calculated between the ratio of the log-ratio (IP/input) of *gcn5-1* over the log-ratio (IP/input) of wild type and the ratio of the log-ratio (IP/input) of *gcn5-2* over the log-ratio (IP/ input) of wild type). Comparative analysis revealed that 89% (2975/3352) of the GCN5 target promoters bound in wild-type plants were precipitated from the chromatin of both mutants, indicating that the truncated GCN5 proteins were still recruited to a large proportion of the target promoters (Figure 3b and Table S1). This result implies that the remaining promoters that were targeted in the wild-type, but not in the mutants, require the GCN5 bromodomain for binding (Figures 2 and 3b).

To assess the reliability of the ChIP–chip results, we used conventional ChIP–PCR to test the enrichment of seven promoters randomly selected from those targeted in the wild-type but not in the *gcn5* mutants. All tested promoters were confirmed to be enriched in the wild-type samples by this independent assay (Figure 3c and Table S2).

Requirement of the GCN5 bromodomain for chromatin binding and acetylation of lysine 14 of histone H3

We have previously shown that the *gcn5-2* mutation results in a clear decrease of histone H3 acetylation in nucleosomes located in the vicinity of the core promoter regions of three light-regulated genes (namely *CAB2*, *RBCS-1A* and *IAA*), and that GCN5 is essential for acetylation of lysine 14 of histone H3 (H3K14; Benhamed *et al.*, 2006). We hypothesized that the bromodomain of GCN5 protein was important for acetylation of H3K14.

To investigate the role of the GCN5 bromodomain in H3K14 acetylation at the genome scale, chromatin of *gcn5-1*, *gcn5-2* and wild-type plants was immunoprecipitated with an antibody that specifically recognizes acetylated H3K14. Our analysis revealed 564 promoters in the mutants for which H3K14 was hypo-acetylated compared to wild-type (Figure 3b). All hypo-acetylated promoters were among the GCN5 binding targets in wild-type plants, indicating a clear functional link between GCN5 binding and H3K14 acetylation. Of these promoters, 215 (38%) were still bound by GCN5 in the mutants. The remaining 349 (62%) promoters corresponded to sites that were no longer bound by the mutant GCN5 proteins, representing 93% (349/377) of the target promoters that required the bromodomain for GCN5 binding. These data suggest that the GCN5 bromodomain is necessary to acetylate H3K14 in a subset of promoter regions. In some cases, lack of acetylation could reflect the inability of the mutant GCN5 protein to bind chromatin. In other cases, the mutant protein was associated with chromatin but unable to acetylate H3K14.

Functional classification of GCN5 targets

Classification of GCN5 target promoters based on gene ontology showed that GCN5 bound to a wide range of promoters of genes with diverse functions (Figure 4), and that GCN5 binding was not obviously biased for any particular gene ontology category. However, promoters bound to

(a)

(b)

(c)
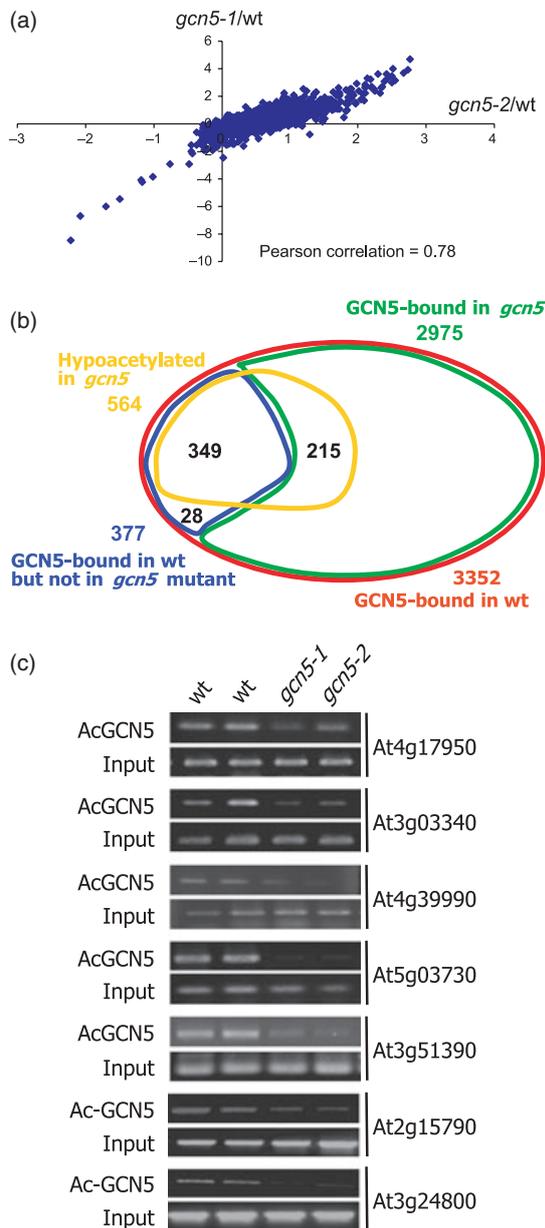
Figure 3. Comparison of promoter targets for GCN5 binding and histone H3K14 acetylation in wild-type, *gcn5-1* and *gcn5-2* plants.

(a) Comparison between ChIP–chip experiments for wild-type (Ws) versus *gcn5-1* or *gcn5-2*.

(b) Venn diagram representing the intersections between the GCN5 promoter targets in the wild-type and *gcn5* mutants and the promoters showing H3K14 hypo-acetylation in *gcn5* mutants. All promoters bound by GCN5 in the wild-type are included in the diagram (3352, shown in red). In the *gcn5* mutants, most promoter targets remained associated with the GCN5 mutant protein (2975, shown in green), while a smaller proportion were no longer bound (377, shown in blue). All promoters hypo-acetylated in *gcn5* mutants versus the wild-type (564, shown in yellow) were bound by GCN5 in the wild-type. Among the promoters that were hypo-acetylated in *gcn5* mutants, 215 remained associated with the GCN5 mutant protein but 349 were no longer bound in the *gcn5* mutants.

(c) PCR confirmation of ChIP products with primers corresponding to seven promoters enriched in the wild-type compared to *gcn5* as determined by ChIP–chip assays.

GCN5 only in the wild-type but not in the *gcn5* mutants, or H3K14 hypo-acetylated in the *gcn5* mutants, preferentially belonged to gene classes coding for proteins involved in transcription (Figure 4 and Table 1).

GCN5 binding versus transcriptional regulation

To determine whether GCN5 binding affected transcriptional regulation, we characterized the transcriptome of wild-type, *gcn5-1* and *gcn5-2* plants by hybridizing CATMA arrays, including 24 576 gene-specific tags (Hilson *et al.*, 2004). Wild-type and mutant plants were grown and harvested under the same conditions as those for the ChIP–chip assays. Using two biological replicates, we found that the two mutants had similar transcript profiles (Pearson correlation = 0.77; calculated for the signal ratios of *gcn5-1*/wild-type versus *gcn5-2*/wild-type; Figure S2). In total, 3.5% (856) of the genes represented by CATMA array features were differentially expressed (Bonferroni *P*-value < 0.05) in both mutants when compared to the wild-type, a percentage close to the 4.0% for *gcn5-1*-de-regulated genes identified previously using the Affymetrix Arabidopsis AG array representing 8200 genes (Vlachonasios *et al.*, 2003). Of the *gcn5*-de-regulated genes, 411 were also present in the SAPv1 array, of which 155 (38%) are GCN5 binding targets. Among these 155 genes, 97 (63%) and 58 (37%) were down and upregulated in the *gcn5* mutants, respectively.

Comparison of GCN5 binding sites with light-regulated gene profiles

gcn5 plants have a light-hyposensitive phenotype that was most pronounced when grown under continuous far-red light, and certain light-inducible genes have reduced expression and histone acetylation in *gcn5* mutants (Benhamed *et al.*, 2006). However, the proportion of promoters corresponding to genes differentially expressed in light or darkness (Ma *et al.*, 2005) was 38% (467 of 1241 promoters present in the SAPv1 array), almost identical to the genome-wide GCN5 association (Figure 5a).

Recent studies have revealed that the HY5 transcriptional activator of light signaling binds to 3894 genes (Lee *et al.*, 2007). Of these genes, 1475 promoters were present on the SAPv1 array, and 559 (38%) were bound by GCN5 (38% of the GCN5 targets). Of the 559 target promoters, 123 are known to be light-regulated (Figure 5a), suggesting that a significant proportion of light-regulated genes depend on binding of both GCN5 and HY5 for expression.

Finally, we compared the GCN5 target promoters with light-regulated gene expression profiles induced by far-red (FR) or red (R) irradiation in a phyA- or phyB-dependent manner (Tepperman *et al.*, 2001, 2004). Genes differentially expressed after 1 and 24 h are referred to as early and late genes, respectively. When combining both time points, the
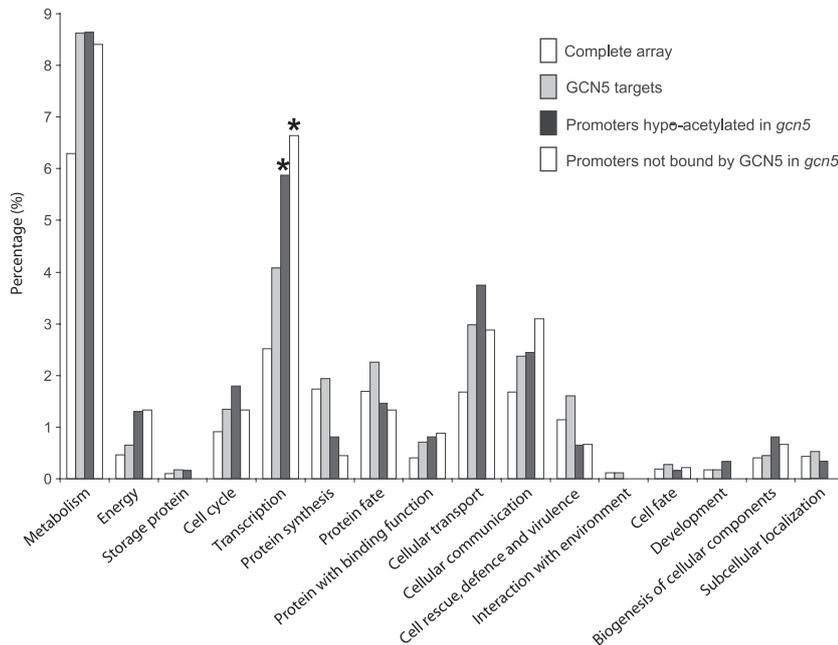
Figure 4. Functional classification of the GCN5 binding target promoters.
Classification was according to the Munich Information Center for Protein Sequences Functional Categories (FunCat). Asterisks indicate significant bias (χ^2 test; $P < 0.05$) in comparison with the subset represented in the SAPv1 array.

proportions of GCN5 target promoters for genes induced or repressed by R/FR irradiation were similar to the genome-wide GCN5 association (Figure 5b and data not shown). However, the proportion of GCN5 target promoters was higher in early induced genes and lower in late induced genes, for both FR and R irradiation (Figure 5b and Table S3). No difference was detected between early and late repressed genes (data not shown), similar to the trends observed for HY5 binding targets (Lee *et al.*, 2007).

Discussion

Multi-purpose Arabidopsis promoter collection

The SAP repertoire designed *in silico* contains 22 256 Arabidopsis promoter sequences that are fully documented in the SAP online database (http://www.psb.ugent.be/SAP) and the FLAGdb++ genome viewer (http://urgv.evry.inra.fr/FLAGdb). Of these, 19 741 were synthesized by PCR from genomic DNA templates. This sequence repertoire is useful to study diverse aspects of transcriptional regulation and genetic networks. Because each SAP promoter PCR amplicon is flanked by *att*B4 and *att*B1 sequences, it can be captured in a plasmid via a Gateway site-specific BP clonase reaction, resulting in a promoter entry clone (Cheo *et al.*, 2004; Dupuy *et al.*, 2004; Sasaki *et al.*, 2004). From such a clone, the promoter sequence can be transferred via an LR clonase reaction into a wide variety of MultiSite Gateway destination vectors, designed for functional assays in plant cells (Karimi *et al.*, 2005, 2007) or in heterologous systems. As a result of the availability of Gateway promoter repertoires, large-scale functional screens have been initiated to dissect the net-

works of transcriptional regulators and their target DNA sequences (Deplancke *et al.*, 2004, 2006) or to systematically record *in vivo* transcriptional activity (Dupuy *et al.*, 2007). Already, over 200 Arabidopsis SAP accessions have been cloned and sequence-verified for downstream applications (unpublished results). In an independent study focusing on genes specifically induced or repressed upon lateral root initiation, transcriptional fusions between each of the 14 tested SAP promoters and the GUS reporter gene confirmed, in transgenic Arabidopsis plants, the pattern deduced from transcript profiles of pericycle sorted cells (Birnbaum *et al.*, 2003; I. De Smet, M. Naudts, R.D., P.H. and T. Beeckman, unpublished results, Department of Plant Systems Biology, Gent, Belgium). This result indicates that, in most cases, promoter sequences of up to 2 kb are sufficient to recapitulate genuine transcriptional patterns. Additionally, the collection of promoter DNA fragments can be spotted onto microarrays for analysis of chromatin modification or localization of factors associated with chromatin *in vivo*.

Genome-wide promoter targets of Arabidopsis GCN5

The promoter targets of the Arabidopsis HAT GCN5 were analyzed by ChIP–chip using the SAPv1 array containing 8471 transcriptionally active promoters. Our data showed that GCN5 targets a large proportion of the promoters in the Arabidopsis genome, distributed homogenously across all five chromosomes, suggesting that GCN5-containing complexes play an essential role in plant transcriptional initiation.

The GCN5 C-terminal bromodomain mediates chromatin association through binding of acetylated histone tails

Table 1 Promoters of transcription factors bound by GCN5 in the wild-type, but not bound and H3K14 hypo-acetylated in the *gcn5* mutants

AGI code	Gene description
At1g06070	bZIP transcription factor, putative (bZIP69)
At1g06850	Basic leucine zipper 52
At1g12980	AP2/ERF protein (ESR1)
At1g14510	PHD finger family protein
At1g52150	Member of the class III HD-ZIP protein family INCURV-ATA 4
At2g03710	MADS box protein AGL3 SEPALLATA4
At2g21320	Zinc finger (B-box type) family protein
At2g23320	WRKY DNA-binding protein 15
At2g26990	Constitutive photomorphogenic 12
At2g30470	HSI2, high-level expression of sugar-inducible 2
At2g35310	Transcriptional factor B3 family protein
At3g12720	Member of the R2R3 factor gene family
At3g27260	DNA-binding bromodomain-containing protein GTE8
At3g28210	Putative zinc finger protein (PMZ)
At3g46130	Putative transcription factor MYB48
At3g53500	RSZ32 nucleic acid binding
At3g54220	Scarecrow transcription factor, putative SGR1
At3g54620	Basic leucine zipper 25
At3g55370	Zinc finger protein OBP3
At3g57670	Zinc finger (C2H2 type) family protein
At3g61950	Basic helix-loop-helix (bHLH) family protein
At3g61970	DNA-binding protein NGA2 (NGATHA2)
At3g62340	WRKY transcription factor WRKY68
At4g07950	DNA-directed RNA polymerase III family protein
At4g14220	Zinc finger (C3HC4-type RING finger) family protein
At4g17950	DNA-binding family protein
At4g29000	Tesmin/TSO1-like CXC domain-containing protein
At4g32800	Member of the DREB subfamily A-4
At4g32850	Poly(A) polymerase IV
At5g02030	BEL1-like homeodomain 8
At5g44080	bZIP transcription factor family protein
At5g46350	WRKY transcription factor
At5g51190	ERF (ethylene response factor) subfamily B-3
At5g59450	Scarecrow-like transcription factor 11 (SCL11)
At5g59780	Putative transcription factor MYB59
At5g63280	Zinc finger (C2H2 type) family protein
At5g67110	Myc/bHLH transcription factor-like protein ALCATRAZ

(Imoberdorf *et al.*, 2006; Mujtaba *et al.*, 2007), but its disruption in two T-DNA insertion mutants abolished binding to only 11% of the GCN5 promoter targets, implying that the bromodomain is not necessary for recruitment to most target promoters in Arabidopsis. Therefore, association with GCN5 probably also relies on additional interactions with DNA-binding transcription factors or with other subunits in various HAT complexes, as direct GCN5–DNA interaction has not been reported to date.

Binding of GCN5 to target promoters does not always correlate with transcriptional regulation

Our data revealed that partial or full deletion of the GCN5 bromodomain resulted in de-regulation of 3.5% of the Arabidopsis genes. Only 38% (155/411) of the de-regulated
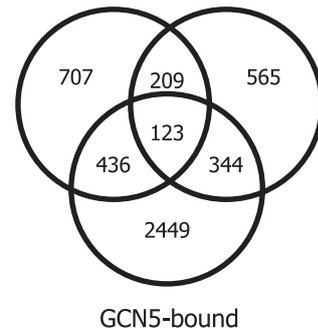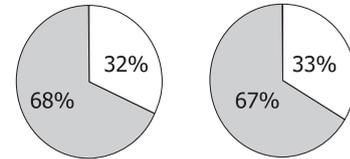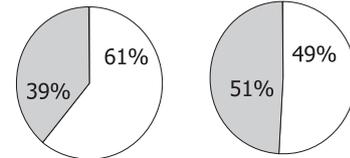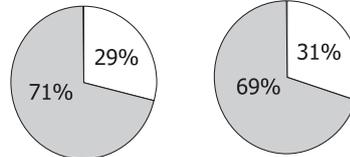
Figure 5. GCN5 binding to the promoters of light-regulated genes and its relationship with the light-signaling transcriptional activator HY5.
(a) Venn diagram comparing light-regulated genes with the binding targets of GCN5 and HY5. Of the light-regulated genes (Ma *et al.*, 2005), 1241 were represented on the SAPv1 array, of which 467 were GCN5 targets. Of the HY5 targets (Lee *et al.*, 2007), 1475 were represented on the SAPv1 array, of which 559 were GCN5 targets, and 123 light-regulated genes were targets of both GCN5 and HY5.
(b) Comparison between fraction of promoters bound by GCN5 for early and late genes induced by FR or R light at an early or late stage. The differences between early- and late-induced GCN5 targets were statistically significant (χ^2 test).

genes overlapped with the GCN5 targets, suggesting that de-regulation of most genes is an indirect consequence of the *gcn5* mutations. For example, the *gcn5* mutations might affect the global histone acetylation level that is required for expression of gene subsets, because the bromodomain might be involved in propagation of chromatin acetylation (Imoberdorf *et al.*, 2006; Mujtaba *et al.*, 2007).

In contrast, *gcn5* mutations did not alter the expression of 62% of the GCN5 target genes. Indeed, GCN5 binding

does not always affect gene transcription, as previously illustrated by its constitutive association with light-responsive promoters irrespective of light conditions (Benhamed *et al.*, 2006). In such cases, specific signaling pathways might regulate GCN5 activity in response to cellular cues. Alternatively, GNC5 may be associated with nucleosomes positioned in chromatin regions that do not regulate the transcriptional activity of neighboring genes, and may be located within or outside the genome fragments included in the SAP promoter amplicons. That *gcn5-1* and *gcn5-2* are not null alleles might also be the reason for the lack of *gcn5* transcriptional de-regulation: both mutants still produce truncated polypeptides (Figure S1), and only the *gcn5-2* mutation removes the entire bromodomain.

Nevertheless, 37% (58/155) of the *gcn5*-de-regulated genes are induced in the mutants (Figure 5), possibly due to indirect effects mediated by the de-regulated transcription factors. Alternatively, GCN5 might exert a transcriptional repressor activity on some genes that could be mediated by acetylation of protein substrates other than histones, such as DNA-binding factors or general factors of the transcription machinery. Finally, HAT activities could also be required for nucleosome assembly during repression, as it is for the developmental regulation of yeast genes (Recht *et al.*, 2006).

Bromodomain involvement in promoter binding and H3K14 acetylation

Bromodomains are evolutionarily conserved proline-rich protein modules that are found mostly in HAT complexes and that interact with the N-terminal tails of histones H3 and H4 with acetylated lysines. The *gcn5* mutations affect acetylation of several histone H3 lysine residues, including K9, K14 and K27, on light-inducible gene promoters (Benhamed *et al.*, 2006). Acetylation of H3K9 and H3K27, however, was also reduced by mutation of *TAF1*, which is another *HAT* gene that is involved in light-controlled transcriptional regulation (Bertrand *et al.*, 2005). Therefore, H3K14 acetylation of the tested light-responsive promoters appears to depend specifically on the intact bromodomain of GCN5. Only 7% (546/8471) of the SAPv1 promoters were H3K14 hypo-acetylated in the *gcn5* mutants, indicating that, for most promoters in the genome, the GCN5 bromodomain is not essential for this modification. Alternatively, the bromodomain might be required for H3K14 acetylation of the ORF during transcriptional elongation, which is mediated by the GCN5-containing SAGA HAT complex (Govind *et al.*, 2007). We cannot exclude the possibility that the GCN5 bromodomain is also involved in global maintenance of chromatin acetylation but not in modification of specific loci. Interestingly, 93% (349/377) of the target promoters that required the bromodomain for GCN5 binding also showed H3K14 hypo-acetylation in the *gcn5* mutants. Thus, for a subset of GCN5 promoter targets, GCN5 recruitment and H3K14 acetylation are coupled and controlled by the GCN5 bromodomain.

GCN5 in light regulation

GCN5 is required for light-regulated gene expression and hypocotyl elongation (Benhamed *et al.*, 2006). We found that the ratio of GCN5-bound versus non-bound targets was similar across all promoters tested (40%) and among light-regulated genes (38%). However, 61% and 49% of the early gene promoters induced by FR and R light, respectively, are GCN5 binding targets. HY5, which has been proposed to function as a high hierarchical regulator of the transcriptional cascade controlling photomorphogenesis, also targets more than 60% of the early FR- and R-induced genes (Lee *et al.*, 2007). Recent studies have shown that acetylation of histones in the yeast *PHO5* gene by a HAT complex under repressed conditions (high phosphate) is essential for the binding of PHO4, a DNA-binding transcription factor that is activated upon phosphate starvation (Nourani *et al.*, 2004). Similarly, GCN5 might prime for later activation of light-inducible genes in the dark repressed state. Because a large proportion of HY5 target promoters induced early by FR or R light were also targets of GCN5 (Table S3), HY5 and GCN5 might cooperate to induce expression of these genes. In support of this hypothesis, genetic interactions between HY5 and GCN5 have been found in the regulation of photomorphogenesis (Benhamed *et al.*, 2006). Whether GCN5 and HY5 interact directly remains to be established.

Experimental procedures

In silico *design of promoter amplicons*

Promoter sequences were extracted from the intergenic regions defined in the genome annotation of the Arabidopsis Information Resource (TAIR version 5). Only sequences upstream of translation initiation codons were considered, thereby discarding intergenic regions between two convergent genes. Genes described as putative transposable elements or parts thereof were also eliminated. When divergent, both promoters derived from the same intergenic region were selected independently and relative to their respective translational start site. For annotated splice variants, only the largest gene model was considered.

In all cases, the 3′ primer defining the promoter amplicon was positioned immediately next to, but excluding, the translation initiation codon, and only its length could vary when searching for a matching PCR primer pair. In most cases, the 5′ primer was positioned close to the next upstream coding sequence when the intergenic region was <2 kb, or approximately 2 kb upstream of the cognate translation initiation codon when the intergenic region was >2 kb. In a minority of cases (1128; 5.5%), the designed promoter amplicons were >2.2 kb and included the entire intergenic region. The 5′ primer was selected within a window centered on the desired boundary and approximately 10% of the target size, thereby enabling some flexibility in the search for matching PCR primers. PCR primers were designed in three successive rounds

using software based on the OSP.pm perl library: (i) primers with size ranging from 18 to 28 nucleotides; (ii) primers with maximum size increased to 32 nucleotides; (iii) primers with melting temperature range relaxed from 55-65 to 50-70°C. The resulting primer pairs were tested for specificity across the whole genome. Each promoter-specific primer sequence (average 28 nucleotides) was extended at the 5′ end by 15 bases, including a trinucleotide specificity tag (TST) and a portion of the *att*B4 and *att*B1 Gateway recombination site in the 5′ and 3′ primer, respectively (Figure 1 and Table S4).

Promoter synthesis

SAPv1 and SAPv2 were initially amplified in a two-step PCR (Figure 1). In the first step, 50 ng of Col-0 genomic DNA and 1 μM of paired specific primers were resuspended in a 10 μl final volume of PCR mix with 1 unit of Platinum *Taq* high-fidelity DNA polymerase (Invitrogen, http://www.invitrogen.com), and incubated for 2 min at 94°C, then for 11 cycles of 30 sec at 94°C, 30 sec at 55°C and 2 min at 68°C, then held at 4°C. In the second step, 40 μl of PCR mix was added to the initial reaction, containing 5 μM of the universal attB4TST and attB1TST oligonucleotides and 1 unit of fresh platinum *Taq* high-fidelity DNA polymerase, and incubated 2 min at 94°C, then for 30 cycles of 30 sec at 94°C, 30 sec at 55°C and 2 min at 68°C, followed by 5 min at 68°C, then held at 4°C. The 40 universal primers carried either the full *att*B4 or *att*B1 Gateway recombination site and the 3′ TST defining each of the 16 rows or 24 columns of a 384-well plate (Table S4). These 40 primers were sufficient for re-amplification of the entire SAP collection. TSTs distinguished rows and columns and helped resolve potential well-to-well cross-contamination. The SAP promoter amplicon collection is available upon request.

Subsequent SAPv1 PCR amplifications were carried out using 2 μl of a threefold dilution in double-distilled water (ddH₂O) of the primary product as template, 2 units of *Taq* DNA polymerase in a 100 μl final volume, and the following program: 2 min at 94°C, 30 cycles of 30 sec at 94°C, 30 sec at 55°C and 2 min at 68°C, then 15 min at 68°C, then held at 4°C. To assess the success of the PCR, 8 μl from each reaction was loaded on a 2% agarose gel, and the size of every amplicon was verified. SAP amplicons were purified on Multiscreen plates (Millipore, http://www.millipore.com) and resuspended in Tris/EDTA/dimethyl sulfoxide at approximately 100 ng ml⁻¹. The purified probes were transferred to 1536-well plates using a Genesis workstation (TECAN, http://www.tecan.com). Although amplification did not yield a detectable product of the expected size in all cases, promoter probes were not rearranged to avoid arraying errors. Each spot was flagged with a PCR product quality code: 0, not detected (3317, 27.9%); 1, expected size (8471, 71.2%); 2, wrong size (116, 1%).

SAP array production

Amplicons were spotted on UltraGAPS slides (Corning, http://www.corning.com) by means of a Microgrid II (Genomic Solutions, http://www.genomicsolutions.com) using a 64-pin tool. The array design is fully documented as array accession number A-MEXP-817 in the ArrayExpress repository (http://www.ebi.ac.uk/arrayexpress/). Pin performance and print quality (spot diameter and homogeneity) were assessed before each SAP DNA printing run. For each batch, quality was controlled on two SAPv1 arrays colored with POPO-3 iodide (Invitrogen). Only arrays with >99% good spots were further processed (same quality threshold as for CATMA arrays).

Plant material

The *gcn5-1* (Vlachonasios *et al.*, 2003) and *gcn5-2* (Bertrand *et al.*, 2003) T-DNA insertion mutant alleles of *Arabidopsis thaliana* (L.) Heyhn. are both in the Wassileskija (Ws) ecotype. Prior to germination, seeds of Ws, *gcn5-1* and *gcn5-2* were sterilized and stratified for 2 days at 4°C. For *in vitro* cultures, seeds were sown on 0.5× Murashige and Skoog medium, incubated at 4°C for 48 h, and transferred to growth chambers for germination at 20°C under white light (16 h at 120 μmol m⁻² sec⁻¹).

Protein gel blots

For protein gel blots, nuclear proteins were extracted from wild-type or *gcn5* plants by grinding rosette leaves to a fine powder in liquid nitrogen. The powder was transferred to nuclear extraction buffer (0.4 M sucrose, 10 mM Tris/HCl pH 8, 10 mM MgCl₂, 5 mM β-mercaptoethanol and 0.1 mM PMSF). The extracts were filtered twice through Miracloth (Calbiochem, http://www.emdbiosciences.com), and centrifuged at 4000 *g* for 20 min at 4°C. The nuclear pellets were dissolved in 15% SDS–PAGE gel-loading buffer for electrophoretic analysis. The blots were incubated with the GCN5 antibody (1:500 dilution) overnight at 4°C, and revealed by enhanced chemiluminescence (GE Healthcare, http://www.gehealthcare.com).

Chromatin immunoprecipitation and SAP array hybridization

Immunoprecipitations were performed as described previously (Gendrel *et al.*, 2005). Twelve-day-old seedlings grown under long-day conditions were harvested at developmental stage 1.03 (Boyes *et al.*, 2001), 4 h after dawn, fixed in 1% formaldehyde for 15 min in a vacuum, and neutralized with 0.125 M glycine. After washing with sterilized water, samples were ground in liquid nitrogen as described above. Nuclear pellets were suspended in a buffer containing 0.25 M sucrose, 10 mM Tris/HCl pH 8, 10 mM MgCl₂, 1% Triton X-100, 5 mM β-mercaptoethanol, 0.1 mM PMSF and protease inhibitors (one mini-tablet per ml; Roche Diagnostics, http://www.roche.com). The suspensions were transferred to microfuge tubes and centrifuged at 12 000 *g* for 10 min. The pellets were suspended in 1.7 M sucrose, 10 mM Tris/HCl pH 8, 2 mM MgCl₂, 0.15% Triton X-100, 5 mM β-mercaptoethanol, 0.1 mM PMSF and protease inhibitors, and centrifuged at 12 000 *g* for 1 h through a layer of the same buffer in microfuge tubes. The nuclear pellets were lysed in a buffer containing 50 mM Tris/HCl pH 8, 10 mM EDTA, 1% SDS and protease inhibitors. The lysed nuclei were sonicated four times for 15 sec at 4°C, followed by centrifugation at 12 000 *g* for 10 min. The supernatants containing chromatin fragments were diluted 10-fold with 1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris/HCl pH 8, and 167 mM NaCl. A 1 ml aliquot of the dilution was used per immunoprecipitation assay. Antibodies against acetyl histone H3K14 and GCN5 were purchased from Upstate Biotechnology (http://www.millipore.com) and custom-prepared in rabbits (Eurogentec, http://www.eurogentec.com; Benhamed *et al.*, 2006), respectively.

Labeling and hybridization were adapted from the methods described by Gendrel *et al.* (2005). DNA from the input chromatin fraction prior to immunoprecipitation (input control) or recovered after ChIP (IP fraction) was differentially labeled and hybridized in classical dye swap experiments to correct for dye bias. The IP and input fractions were precipitated and amplified as follows. The first step was carried out in 16.5 μl final volume: 2 min at 94°C, 10 min at 5°C, addition of the mix containing primer A (5′-GTTTCCCAGTCAC-GATCNNNNNNNNNN-3′), 8 min at 37°C, 2 min at 94°C, 5 min at 10°C,

addition of 4 μl Sequenase (USB Corporation, http://www.usbweb.com), 8 min at 37°C, then held at 4°C. After addition of 43.5 μl ddH$_2$O, 15 μl of the first reaction dilution were used as template for the amplification step, after addition of primer B (5′-GTTTCCCAGTCACGATC-3′), in 100 μl final volume, with the following program: 3 min hold at 94°C; 30 cycles of 30 sec at 94°C, 30 sec at 40°C, 30 sec at 50°C and 1 min at 72°C; then held at 4°C.

PCR products were purified using Qiaquick columns (Qiagen, http://www.qiagen.com) according to the manufacturer's instructions, and labeled with either Cy3-dCTP or Cy5- dCTP (NEN, las.perkinelmer.com) using 10 units of DNA Klenow polymerase and random hexamers. Labeled samples were purified using Qiaquick columns, quantified, combined and precipitated. At least 40 pmol of each labeled sample was used per hybridization. Slides were pre-hybridized for 1 h, hybridized overnight at 42°C in 25% formamide, washed in 2× SSC, 0.1% SDS for 4 min at 42°C, and in 1× SSC for 4 min, 0.2× SSC for 4 min and 0.05× SSC for 4 min at room temperature, and dried by centrifugation at 1000 *g* for 2 min. A technical replicate in the dye swap was used for each comparison. The arrays were scanned using a GenePix 4000A scanner (Axon Instruments, http://www.moleculardevices.com), images were analyzed with the software GenePix Pro 3.0 (Axon Instruments), and data were stored as a GenePix reader (gpr) file. The ChIP–chip data have been submitted to Array Express under accession number E-MEXP-1233.

ChIP–chip SAPv1 array data analysis

Pre-processing data and normalization were as described previously (Turck *et al.*, 2007). Raw data (log$_2$ of the median feature pixel intensity as listed in gpr files, without background subtraction) for the 635 and 532 nm signals were processed, excluding spots manually flagged because of sub-standard features. Array-by-array normalization, such as loess, could not be applied because the signal distribution of the input and IP samples differed substantially. Instead, data were normalized across arrays based on the properties of the dye swap. As the second array was a technical replicate of the first one, the spot signal of the sample hybridized on the first array and labeled with the red dye (or the green dye) should be close to the green signal of the second array (or the red signal, respectively). In practice, the relationship is linear, but it is not the identity function. To correct this gap, we used these following linear models: $Y_{21s} = a + bY_{12s} + N(0, \sigma^2)$ and $Y_{22s} = c + dY_{11s} + N(0, \sigma^2)$, where Y_{ijs} is the signal of the spot 's' when the sample hybridized to array 'i' is labeled with dye 'j', and the normalized IP and input values of spot 's' on the second array relative to the first one are defined by $Y_{21s} = (Y_{21s} − a)/b$ and $Y_{22s} = (Y_{22s} − c)/d$, respectively. Then the values of the two arrays in the same dye-swap were averaged for each sample and for each spot.

For each dye swap, 'enriched' probes were defined as probes yielding a signal significantly higher than the background. To this end, the IP signal histogram was modeled according to mixture models. Truncated and non-truncated Gaussian mixture models were applied, with a component number varying from 1 to 5. Models with two or three components were systematically selected using the Bayesian information criterion. The components were interpreted as determining significant IP signals. The selected model was often a mixture with two well separated components. In such a case, probes classified in the component with the greatest mean according to the maximum a posteriori (MAP) rule were considered to yield a significant IP value. When the number of components was three, the component with the lowest mean was well separated from the others and characterized background hybridization. Therefore, probes classified in the other two compo-

nents according to the MAP rule were considered to yield a significant IP value. A promoter is considered bound when it yields significant IP values in the three biological replicates.

Transcriptome CATMA array hybridization

Transcriptome profiling was analyzed using Arabidopsis CATMA arrays containing 24 576 gene-specific tags (CATMAv2 repertoire; Crowe *et al.*, 2003; Hilson *et al.*, 2004). Arrays were printed and hybridized as described previously (Lurin *et al.*, 2004). RNA was extracted from 12-day-old seedlings at developmental stage 1.03 grown under long-day conditions (Boyes *et al.*, 2001). Each experiment was carried out with two independent biological repeats to take into account biological variation. Each comparison included a technical repeat with fluorochrome reversal (dye swap) for each pool of RNA. RNA integrity was verified using a Bioanalyzer (Agilent, http://www.agilent.com). For each sample, cRNAs were synthesized from 2 μg of total RNA using the Message Amp aRNA kit (Ambion, http://www.ambion.com), and 5 μg of cRNA was reverse-transcribed using 300 units of SuperScript II (Invitrogen) and Cy3-dUTP or Cy5-dUTP (NEN) for each slide. Labeled samples were combined, purified, and concentrated using YM30 Microcon columns (Millipore). The CATMA arrays were pre-hybridized, hybridized, washed, dried and scanned as for SAPv1 arrays. The transcript profiling data have been submitted to Array Express under accession number E-MEXP-1241.

CATMA array data analysis

Transcriptome statistical analysis was based on dye swaps (i.e. two arrays each containing 24 576 GSTs and 384 controls). For each array, the raw data consisted of the log$_2$ median feature pixel intensity without background subtraction, at 635 nm (red) and 532 nm (green). Data analysis was performed as described previously (Gagnot *et al.*, 2008). In the following description, log ratio refers to the differential expression between two conditions: either log$_2$ red/green or log$_2$ green/red according to the experimental design. Array-by-array normalization was performed to remove systematic biases: spots with poor features were excluded; to correct for dye bias, signal was normalized globally according to intensity using the LOESS procedure; and, for each block, the log ratio median calculated over the values for the entire block was subtracted from each individual log ratio value to correct for print tip effects in each metablock. To determine differentially expressed genes, the log ratios were analyzed using a paired *t*-test, assuming that the variance of the log ratios is the same for all genes. Spots with extreme variance (high or low) were excluded. The raw *P* values were adjusted by the Bonferroni method (type I error equal to 5%) to control for family-wise error rate and to limit false-positives drastically in a multiple comparison context (Benjamini and Hochberg, 1995). The data were deposited in the ArrayExpress (http://www.ebi.ac.uk/arrayexpress/) according to MIAME standards.

Acknowledgements

Detlef Weigel, Max Planck Institute for Developmental Biology, Tübingen, Germany) for generation of the promoter resource. We thank Martine De Cock for help in preparing the manuscript and Karel Spruyt for assistance with illustrations.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Protein gel analysis of GCN5 proteins from whole-cell extracts of the wild-type (Ws) and *gcn5* mutants.

Figure S2. Comparison of transcript profiles between *gcn5-1*, *gcn5-2* and the wild-type.

Table S1. Transcripts and chromatin targets that differ between the wild-type and *gcn5* mutants.

Table S2. Primers for PCR analysis of GCN5 chromatin immuno-precipitation.

Table S3. GCN5 and HY5 binding with respect to light transcriptional regulation.

Table S4. Universal primers for SAP promoter secondary amplification.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Benhamed, M., Bertrand, C., Servet, C. and Zhou, D.-X. (2006) *Arabidopsis GCN5, HD1,* and *TAF1/HAF2* interact to regulate histone acetylation required for light-responsive gene expression. *Plant Cell*, **18**, 2893–2903.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R Stat. Soc. B*, **57**, 289–300.

Bertrand, C., Bergounioux, C., Domenichini, S., Delarue, M. and Zhou, D.-X. (2003) Arabidopsis histone acetyltransferase AtGCN5 regulates the floral meristem activity through the *WUSCHEL/AGAMOUS* pathway. *J. Biol. Chem.* **278**, 28246–28251.

Bertrand, C., Benhamed, M., Li, Y.-F., Ayadi, M., Lemonnier, G., Renou, J.-P., Delarue, M. and Zhou, D.-X. (2005) *Arabidopsis* HAF2 gene encoding TATA-binding protein (TBP)-associated factor TAF1, is required to integrate light signals to regulate gene expression and growth. *J. Biol. Chem.* **280**, 1465–1473.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W. and Benfey, P.N. (2003) A gene expression map of the *Arabidopsis* root. *Science*, **302**, 1956–1960.

Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R. and Görlach, J. (2001) Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell*, **13**, 1499–1510.

Carré, C., Szymczak, D., Pidoux, J. and Antoniewski, C. (2005) The histone H3 acetylase dGcn5 is a key player in *Drosophila melanogaster* metamorphosis. *Mol. Cell Biol.* **25**, 8228–8238.

Cheo, D.L., Titus, S.A., Byrd, D.R.N., Hartley, J.L., Temple, G.F. and Brasch, M.A. (2004) Concerted assembly and cloning of multiple DNA segments using in vitro site-specific recombination: functional analysis of multi-segment expression clones. *Genome Res.* **14**, 2111–2120.

Cosma, M.P., Tanaka, T. and Nasmyth, K. (1999) Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell*, **97**, 299–311.

Crowe, M.L., Serizet, C., Thareau, V. *et al.* (2003) CATMA: a complete *Arabidopsis* GST database. *Nucleic Acids Res.* **31**, 156–158.

Deplancke, B., Dupuy, D., Vidal, M. and Walhout, A.J.M. (2004) A Gateway-compatible yeast one-hybrid system. *Genome Res.* **14**, 2093–2101.

Deplancke, B., Mukhopadhyay, A., Ao, W. *et al.* (2006) A gene-centered *C. elegans* protein–DNA interaction network. *Cell*, **125**, 1193–1205.

Dupuy, D., Li, Q.-R., Deplancke, B. *et al.* (2004) A first version of the *Caenorhabditis elegans* promoterome. *Genome Res.* **14**, 2169–2175.

Dupuy, D., Bertin, N., Hidalgo, C.A. *et al.* (2007) Genome-scale analysis of in vivo spatiotemporal promoter activity in *Caenorhabditis elegans*. *Nat. Biotechnol.* **25**, 663–668.

Gagnot, S., Tamby, J.P., Martin-Magniette, M.L., Bitton, F., Taconnat, L., Balzergue, S., Aubourg, S., Renou, J.P., Lecharny, A. and Brunaud, V. (2008) CATdb: a public access to Arabidopsis transcriptome data from the URGV-CATMA. *Nucleic Acids Res.* **36**, D986–D990.

Gendrel, A.-V., Lippman, Z., Martienssen, R. and Colot, V. (2005) Profiling histone modification patterns in plants using genomic tiling microarrays. *Nat. Methods*, **2**, 213–218.

Govind, C.K., Zhang, F., Qiu, H., Hofmeyer, K. and Hinnebusch, A.G. (2007) Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. *Mol. Cell*, **25**, 31–42.

Hanlon, S.E. and Lieb, J.D. (2004) Progress and challenges in profiling the dynamics of chromatin and transcription factor binding with DNA microarrays. *Curr. Opin. Genet. Dev.* **14**, 697–705.

Hilson, P., Allemeersch, J., Altmann, T. *et al.* (2004) Versatile gene-specific sequence tags for *Arabidopsis* functional genomics: transcript profiling and reverse genetics applications. *Genome Res.* **14**, 2176–2189.

Horn, P.J. and Peterson, C.L. (2002) Chromatin higher order folding: wrapping up transcription. *Science*, **297**, 1824–1827.

Imoberdorf, R.M., Topalidou, I. and Strubin, M. (2006) A role for Gcn5-mediated global histone acetylation in transcriptional regulation. *Mol. Cell. Biol.* **26**, 1610–1616.

Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science*, **293**, 1074–1080.

Karimi, M., De Meyer, B. and Hilson, P. (2005) Modular cloning in plant cells. *Trends Plant Sci.* **10**, 103–105.

Karimi, M., Bleys, A., Vanderhaeghen, R. and Hilson, P. (2007) Building blocks for plant gene assembly. *Plant Physiol.* **145**, 1183–1191.

Lee, K.K. and Workman, J.L. (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nat. Rev. Mol. Cell Biol.* **8**, 284–295.

Lee, T.I., Causton, H.C., Holstege, F.C.P., Shen, W.-C., Hannett, N., Jennings, E.G., Winston, F., Green, M.R. and Young, R.A. (2000) Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature*, **405**, 701–704.

Lee, J., He, K., Stolc, V., Lee, H., Figueroa, P., Gao, Y., Tongprasit, W., Zhao, H., Lee, I. and Deng, X.W. (2007) Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. *Plant Cell*, **19**, 731–749.

Long, J.A., Ohno, C., Smith, Z.R. and Meyerowitz, E.M. (2006) TOPLESS regulates apical embryonic fate in *Arabidopsis*. *Science*, **312**, 1520–1523.

Lurin, C., Andrés, C., Aubourg, S. *et al.* (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089–2103.

Ma, L., Sun, N., Liu, X., Jiao, Y., Zhao, H. and Deng, X.W. (2005) Organ-specific expression of Arabidopsis genome during development. *Plant Physiol.* **138**, 80–91.

Mao, Y., Pavangadkar, K.A., Thomashow, M.F. and Triezenberg, S.J. (2006) Physical and functional interactions of *Arabidopsis* ADA2 transcriptional coactivator proteins with the acetyltransferase GCN5 and with the cold-induced transcription factor CBF1. *Biochim. Biophys. Acta*, **1759**, 69–79.

Millar, C.B. and Grunstein, M. (2006) Genome-wide patterns of histone modifications in yeast. *Nat. Rev. Mol. Cell Biol.* **7**, 657–666.

Mujtaba, S., Zeng, L. and Zhou, M.-M. (2007) Structure and acetyl-lysine recognition of the bromodomain. *Oncogene*, **26**, 5521–5527.

Nourani, A., Utley, R.T., Allard, S. and Côté, J. (2004) Recruitment of the NuA4 complex poises the *PHO5* promoter for chromatin remodeling and activation. *EMBO J.* **23**, 2597–2607.

Pandey, R., Müller, A., Napoli, C.A., Selinger, D.A., Pikaard, C.S., Richards, E.J., Bender, J., Mount, D.W. and Jorgensen, R.A. (2002) Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Res.* **30**, 5036–5055.

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., III and Grant, P.A. (2005) Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature*, **433**, 434–438.

Recht, J., Tsubota, T., Tanny, J.C. et al. (2006) Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc. Natl Acad. Sci. USA*, **103**, 6988–6993.

Robert, F., Pokholok, D.K., Hannett, N.M., Rinaldi, N.J., Chandy, M., Rolfe, A., Workman, J.L., Gifford, D.K. and Young, R.A. (2004) Global position and recruitment of HATs and HDACs in the yeast genome. *Mol. Cell*, **16**, 199–209.

Samson, F., Brunaud, V., Duchêne, S., De Oliveira, Y., Caboche, M., Lecharny, A. and Aubourg, S. (2004) FLAGdb++: a database for the functional analysis of the *Arabidopsis* genome. *Nucleic Acids Res.* **32**, D347–D350.

Sasaki, Y., Sone, T., Yoshida, S., Yahata, K., Hotta, J., Chesnut, J.D., Honda, T. and Imamoto, F. (2004) Evidence for high specificity and efficiency of multiple recombination signals in mixed DNA cloning by the multisite Gateway system. *J. Biotechnol.* **107**, 233–243.

Tepperman, J.M., Zhu, T., Chang, H.-S., Wang, X. and Quail, P.H. (2001) Multiple transcription-factor genes are early targets of phytochrome A signaling. *Proc. Natl Acad. Sci. USA*, **98**, 9437–9442.

Tepperman, J.M., Hudson, M.E., Khanna, R., Zhu, T., Chang, S.H., Wang, X. and Quail, P.H. (2004) Expression profiling of *phyB* mutant demonstrates substantial contribution of other phytochromes to red-light-regulated gene expression during seedling de-etiolation. *Plant J.* **38**, 725–739.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.-L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G. and Colot, V. (2007) Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet.* **3**, e86, 0855–0866.

Vlachonasios, K.E., Thomashow, M.F. and Triezenberg, S.J. (2003) Disruption mutations of *ADA2b* and *GCN5* transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression. *Plant Cell*, **15**, 626–638.

Yamada, K., Lim, J., Dale, J.M. et al. (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science*, **302**, 842–846.

Yamauchi, T., Yamauchi, J., Kuwata, T., Tamura, T., Yamashita, T., Bae, N., Westphal, H., Ozato, K. and Nakatani, Y. (2000) Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. *Proc. Natl Acad. Sci. USA*, **97**, 11303–11306.

Zhang, X., Yazaki, J., Sundaresan, A. et al. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell*, **126**, 1189–1201.