New Phytologist

Genome-wide search and functional identification of transcription factors in the mycorrhizal fungus *Tuber melanosporum*

Barbara Montanini[1], Elisabetta Levati[1], Angelo Bolchi[1], Annegret Kohler[2], Emmanuelle Morin[2], Emilie Tisserant[2], Francis Martin[2] and Simone Ottonello[1]

[1]Department of Biochemistry and Molecular Biology, University of Parma, 43100 Parma, Italy; [2]Ecogenomics of Interactions Lab, UMR 'Interactions Arbres/Micro-Organismes', INRA-Nancy, 54280 Champenoux, France

Author for correspondence:
Simone Ottonello
Tel: +39 0521 905646
Email: s.ottonello@unipr.it

Received: *6 July 2010*
Accepted: *22 September 2010*

New Phytologist (2011) **189**: 736–750
doi: 10.1111/j.1469-8137.2010.03525.x

Key words: mycorrhiza, regulome, transcription factors, transcriptional activator trap assay, truffles, tuber, yeast.

Summary

- Developmental transitions associated with the life cycle of plant-symbiotic fungi, such as the ascomycete *Tuber melanosporum*, are likely to require an extensive reprogramming of gene expression brought about by transcription factors (TFs). To date, little is known about the transcriptome alterations that accompany developmental shifts associated with symbiosis or fruiting body formation.
- Taking advantage of the black truffle genome sequence, we used a bioinformatic approach, coupled with functional analysis in yeast and transcriptome profiling, to identify and catalogue *T. melanosporum* TFs, the so-called 'regulome'.
- The *T. melanosporum* regulome contains 102 homologs of previously characterized TFs, 57 homologs of hypothetical TFs, and 42 putative TFs apparently unique to *Tuber*. The yeast screen allowed the functional discovery of four TFs and the validation of about one-fifth of the *in silico* predicted TFs. Truffle proteins apparently unrelated to transcription were also identified as potential transcriptional regulators, together with a number of plant TFs.
- Twenty-nine TFs, some of which associated with particular developmental stages, were found to be up-regulated in ECMs or fruiting bodies. About one-quarter of these up-regulated TFs are expressed at surprisingly high levels, thus pointing to a striking functional specialization of the different stages of the *Tuber* life cycle.

Introduction

The life cycle of the symbiotic ascomycete, and prized truffle producer, *Tuber melanosporum* comprises three main stages: free-living mycelium (FLM), ectomycorrhiza (ECM), and fruiting body (FB; 'truffle'). Little is known about the molecular mechanisms underlying morphogenetic transitions between the above life cycle stages, and most of the available information relies on transcript profiling. As revealed by pregenomic cDNA array analyses conducted on basidiomycete ectomycorrhizal associations, mycorrhizal development is indeed accompanied by changes in the abundance of particular groups of transcripts. Genes involved in fungal cell division and proliferation, differentiation and signaling, synthesis of cell wall and extracellular matrix, plant defense or stress response, and primary metabolism appeared to be specifically up-regulated in mycorrhizas (Martin *et al.*, 2007). Similar large-scale gene expression analyses have been conducted in the ectomycorrhizal ascomycetes *Terfezia boudieri* (Zaretsky *et al.*, 2006) and in the whitish truffle *Tuber borchii* (Lacourt *et al.*, 2002; Gabella *et al.*, 2005; Montanini *et al.*, 2006a). In the case of *T. borchii*, nitrogen starvation, an abiotic stress that accompanies (and perhaps contributes to) symbiosis establishment, was used to identify changes in gene expression in stages immediately before mycorrhiza formation. Genes directly involved in nitrogen metabolism as well as genes involved in cell wall and hyphal morphology modification emerged as candidates from the latter analysis, while unknown genes were the main outcome of the study in *T. boudieri* (Zaretsky *et al.*, 2006). By comparison, most genes identified as differentially expressed during FB formation in *T. borchii* were related to cell wall

biosynthesis, antioxidant defense, and nitrogen/carbon metabolism, especially gluconeogenesis and the glyoxylate cycle (Lacourt *et al.*, 2002; Gabella *et al.*, 2005). A more comprehensive picture of gene expression pattern variations associated with specific life cycle stages of ectomycorrhizal fungi has been obtained recently, following completion of the genome sequences of the ectomycorrhizal basidiomycete *Laccaria bicolor* and *T. melanosporum*, the first sequenced plant-symbiotic ascomycete (Martin *et al.*, 2008, 2010). Transcriptome analyses carried out in both studies confirmed, and further detailed, gene expression pattern variations accompanying symbiosis establishment, but also highlighted profound differences in the so-called 'symbiosis toolbox' utilized by these fungi to implement the mycorrhizal program. Furthermore, despite genome-wide expression studies of increasing resolution, transcription factors potentially responsible for the observed changes in gene expression have not been systematically examined in a plant-symbiotic fungus so far.

Ascomycetes are the eukaryotic organisms with the highest genome sequence coverage to date. This, together with the recently released genome sequence of *T. melanosporum* (Martin *et al.*, 2010), allowed access, for the first time, to the entire repertoire of transcription factors (TFs; the so-called 'regulome') in this organism. TFs are regulatory proteins that control gene expression and are involved in a variety of cellular functions. Their basic domain architecture includes at least one DNA-binding domain (DBD) that binds to sequence-specific DNA elements in the promoter region of target genes, and an activation domain (AD) that interacts with (and recruits) the transcription machinery. Additional domains can mediate homodimerization, interaction with other TFs, or the binding of co-activators/co-repressors or small molecule ligands. TFs are classified into superclasses and classes according to their DBDs, which usually display fairly conserved folds and structures (Stegmaier *et al.*, 2004). ADs, instead, are generally more idiosyncratic, lack common structural elements and are usually less easily recognizable. *In silico* TF predictions thus need to be corroborated and extended by functional analyses.

One of such analyses takes advantage of the distinct and independent functions played by TF domains and relies on a modified version of the yeast two-hybrid system, named 'transcriptional activator trap' (TAT) assay, in which selected open reading frames or an entire cDNA library from the organism (and tissue/stage) of interest are cloned downstream to a vector-provided DBD (yGAL4-DBD) (Ye *et al.*, 2004; Titz *et al.*, 2006). When the tested sequence codes for a transcriptional activator, the expression of three reporter genes is activated. Despite the high-throughput capabilities and apparent simplicity of the TAT assay, both false-negative and false-positive results can be expected. The latter, which are perhaps more relevant for the present study, may be caused by forced nuclear internalization of otherwise cytoplasmic proteins as a result of the presence of a vector-provided nuclear localization signal (NLS) associated with the yGAL4-DBD. It should be noted, however, that many instances of apparently cytoplasmic proteins capable of entering the nucleus and activating transcription have been reported recently. These include, for example, actin-related proteins and proteins involved in endocytosis (Pilecka *et al.*, 2007; Zheng *et al.*, 2009), but also a number of metabolic enzymes and other proteins (Hall *et al.*, 2004; Shi & Shi, 2004; Hu *et al.*, 2009), which despite the lack of a recognizable DBD have been found to be capable of DNA binding and transcriptional activation.

The aim of this work was to use state-of-the-art prediction tools in combination with the TAT assay (outlined in Fig. 1) to obtain the largest possible (and less biased) repertoire of DBD-containing proteins and *bona fide* transcriptional regulators encoded by the *T. melanosporum* genome, especially TFs potentially involved in ECM and FB formation. Gene expression programs underlying mycorrhiza formation and symbiosis establishment are still largely unknown. Fruiting body development has been studied extensively in *Neurospora* and *Aspergillus*, but at variance with these model ascomycetes, *T. melanosporum* is a macrofungus producing large-sized underground ascomata, whose formation strictly depends on prior interaction with a host plant. There is thus much potential for the discovery of new TFs, or homologs of previously known TFs reallocated to a new function in *T. melanosporum*. Therefore, in our analysis, special attention was paid to orphan genes, to duplicated genes that may reflect a novel (nonredundant) function in truffles, and to TF genes preferentially expressed in mycorrhiza or fruiting bodies.

Materials and Methods

Bioinformatic and transcriptome analyses

Sequences corresponding to functionally characterized or predicted TFs from *Saccharomyces cerevisiae*, *Neurospora crassa* and *Aspergillus nidulans* (Borkovich *et al.*, 2004; Beskow & Wright, 2006; Colot *et al.*, 2006; Wilson *et al.*, 2008) were employed as queries for a BLASTP search against the predicted *T. melanosporum* proteome, which was conducted at the Génoscope and INRA *Tuber* sites (http://mycor.nancy.inra.fr/IMGC/TuberGenome/, https://www.genoscope.cns.fr/secure-nda/Tuber/html/entry_ggb.html) using a threshold *e*-value of 10^{-5}. This initial set of putative TFs was further scrutinized by a BLASTP search run against the GenBank database in order to remove TF-unrelated as well as mispredicted sequences. All putative proteins containing a recognizable DBD (pfam) were also included in the analysis. Gene products sharing a significant similarity with TFs, but lacking a typical DBD, were individually verified (and

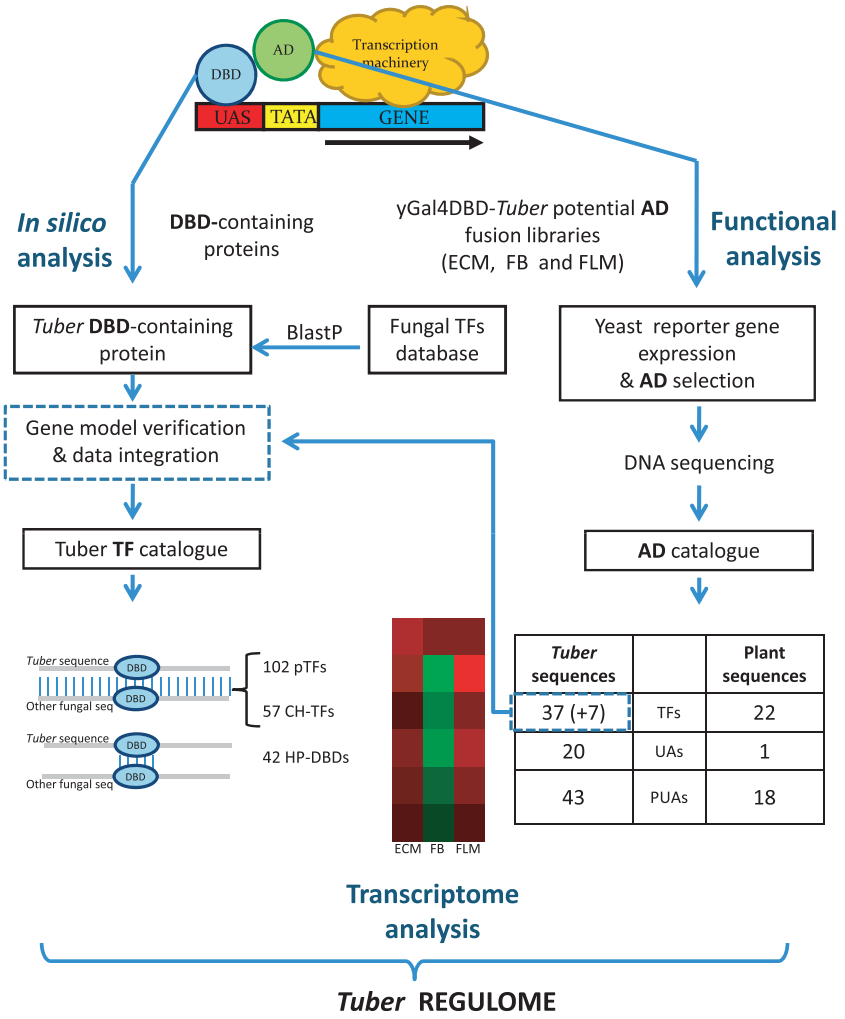

Fig. 1 Outline of the combined, bioinformatic and functional transcription factor (TF) identification strategy utilized in this study. *In silico* analysis, focused on the conserved DNA-binding domain (DBD) region, allowed the assignment of putative *Tuber melanosporum* TFs to different categories, based on sequence similarity with characterized TFs from other fungi (see text and Supporting Information, Table S1 for further details). Heterologous functional analysis in yeast ('transcriptional activator trap', TAT; see text and Fig. S1 for details) was carried out in parallel and led to the validation of 44 *in silico* identified TFs and to the identification of 63 DBD-lacking unconventional activators (UAs and putative UAs (PUAs); see Tables 2, S3). Forty-one TAT-positive proteins of plant origin were retrieved from the screening of the ectomycorrhiza (ECM) library (see Table S2). These analyses, coupled with transcriptome data, led to a comprehensive list of positively acting *Tuber* regulators ('regulome'). pTFs, putative transcription factors; CH-TFs, conserved hypothetical transcription factors; HP-DBDs, DBD-containing hypothetical proteins; FB, fruiting body; FLM, free-living mycelium; AD, activation domain (see text for further details on these definitions).

corrected if necessary) in order to uncover a potentially mispredicted DBD. If a DBD could not be detected, no putative transcription factor function was assigned to the protein.

This set of *Tuber* TFs was then subjected to BLASTP analysis against the Swiss-Prot database, masking PROSITE profiles or Pfam hidden Markov models matching regions in order to distinguish proteins with homology restricted to the DBD from those sharing homology with *bona fide* TFs along the entire sequence. Based on these results, putative *Tuber* TFs were classified as 'DBD-containing hypothetical proteins' (HP-DBD) when sequence homology was restricted to the DBD, or as 'conserved hypothetical TFs' (CH-TF) in the case of proteins for which sequence similarity was also significant outside of the DBD. CH-TFs similar to functionally identified TFs from other fungi (i.e. *Tuber* proteins that are best reciprocal hits of characterized proteins) were separately grouped and designated as 'putative TFs' (pTFs). Phylogenetic analysis (not shown) was used to define paralogous relationships between *Tuber* genes.

Whole-genome exon oligoarray and RNA-Seq analyses were conducted as described (Martin *et al.*, 2010).

Yeast-expressible cDNA library construction

Total RNA (500 μg per sample) was prepared from *T. melanosporum* fruiting bodies and from free-living mycelia (strain Mel28) grown in liquid malt culture for 3 months (kindly provided by Gérard Chevalier, INRA-Clermont-Ferrand) as described elsewhere (Balestrini *et al.*, 2000). The RNeasy Plant Mini Kit (Qiagen) was used for RNA cleaning and DNase I treatment, followed by mRNA purification on Oligotex columns (Qiagen). FB and FLM cDNA libraries were constructed with the CloneMiner cDNA Library Construction Kit (Invitrogen), starting from 2 μg of purified mRNA for each library. For the ECM library, total RNA was extracted from *T. melanosporum*-inoculated hazelnut (*Corylus avellana*) roots (kindly provided by Raffaella Balestrini, University of Torino), with the RNeasy Plant Mini Kit (Qiagen), using RLC buffer

containing 20 mg ml^{-1} PEG 8000 and applying DNaseI on column. mRNA was then retrotranscribed, amplified and cloned, followed by Gateway®-mediated transfer of the initial libraries to the yeast-expressible pDEST™32 vector (supplied with the ProQuest™ Two-Hybrid System kit, Invitrogen).

Functional analysis in yeast

For the 'transcriptional activator trap' assay, yeast strain MaV103 (kindly provided by Marian Walhout, University of Massachusetts) harboring Gal4-dependent *LacZ*, *HIS3* and *URA3* reporter genes was transformed with the three pDEST™32 cDNA libraries and plated on selective medium (SD-Leu-His) containing 25 mM 3-amino-1,2,4-triazole (3AT). At least 500 000 colonies were screened for each library. 3AT-resistant colonies were individually transferred to 384-well SD-Leu plates, handled with a 384-multipinner device (V&P) and analyzed for reporter gene expression. Briefly, for the *HIS3* reporter assay, freshly grown colonies were replica-plated on SD-Leu (positive control), SD-Leu-His + 25 mM 3AT and SD-Leu-His + 100 mM 3AT (test); growth was monitored by visual inspection after 2 and 4 d at 30°C for 25 and 100 mM 3AT, respectively, and colony sizes were compared with those of the corresponding colonies grown for 2 d on positive control plates. Clones yielding large or small colonies on 100 mM 3AT were classified as strong (+++) or medium-strength (++) activators; clones that failed to grow on 100 mM 3AT but which grew on 25 mM 3AT were classified as weak activators (+), whereas clones that did not grow on either medium were considered as false positives (−). An example of this assay is shown in Supporting Information, Fig. S1. For the *URA3* reporter, freshly grown colonies were replica-plated on SD-Leu (positive control) and SD-Leu -Ura (test) for 5 d. After culture, activator strength was classified as none (−), weak (+), medium (++) or strong (+++), based on visual inspection of colony size using the positive control as a reference. For the *LacZ* (β-Gal) reporter assay, candidate clones were pinned onto 150 mm-YPD (1% yeast extract, 2% peptone, 2% dextrose, 2% agar) plates overlaid by a nitrocellulose or nylon membrane. The β-Gal assay was perform as described by Walhout & Vidal (2001). After culture incubation at 37°C for 24 h, clones were classified based on color development as no activation (−, white), weak (+, green), medium (++, light-blue), and strong (+++, dark-blue) (Fig. S1). Cumulative scores from the three assays were used to evaluate activation strength: clones with score values ≥ ++ in at least two assays and ≥ + in the remaining assay were considered as putative strong activators. Despite this fairly stringent criterion, *c.* 70% of the clones that passed initial selection were retained and subjected to sequence analysis.

Targeted functional analysis of individual putative TFs was conducted by cloning each gene in a modified version of the pDEST™32 plasmid, in which the Gateway® recombination cassette is replaced by a *Cpo*I restriction site. *Cpo*I restriction sites flanking the coding sequence of the gene of interest and in frame with the yGAL4-DBD were introduced by PCR. Amplifications were carried out using as template 200 ng of the library corresponding to the life cycle stage in which the TF of interest was most expressed. 3′-rapid amplification of cDNA ends (3′-RACE) was used to identify the C-terminal portion of the protein encoded by GSTUMT00012451001.

Transcriptional activator trap-positive clones were then sequenced, and the resulting sequences were first trimmed to remove low-quality and vector-derived regions. Polished sequences were subsequently assembled into contigs using the EGAssembler (Masoudi-Nejad *et al.*, 2006). The resulting unisequences were used as queries for a BLASTX search against the *Tuber* proteome and, when available, a gene model was assigned to each sequence (identity > 95%). Sequences overlapping the same gene model were checked manually and included in contig sequences. Nucleotide and amino acid sequences were then subjected to BLASTX and BLASTP searches at the National Center for Biotechnology Information. Based on the outcome of these searches, sequences were classified as belonging to either plant or fungi, and to specific structural/functional categories (see the Results section). Sequences without a significant protein match (*e*-value < 10^{-4}) were analyzed with BLASTN. No sequence similar to noncoding RNAs (or vector DNA) was retrieved from the TAT assay. Putative NLSs were searched with the PSORT II prediction tool (Horton & Nakai, 1997).

Results

Predicted *T. melanosporum* regulome

Putative transcription factors bearing a recognizable DBD were identified by searching the *T. melanosporum* genome sequence with BLASTP, using functionally characterized fungal TFs and predicted TFs from *S. cerevisiae*, *N. crassa* and *A. nidulans* as queries (Borkovich *et al.*, 2004; Beskow & Wright, 2006; Colot *et al.*, 2006; Wilson *et al.*, 2008). The resulting candidate sequences were classified according to the TF classification system of Stegmaier *et al.* (2004), taking into account the major classes of fungal TFs: basic domain leucine zipper (bZIP), basic helix-loop-helix (bHLH), and the zinc-coordinating GATA Zn finger and C_2H_2 Zn finger, and the fungal Zn_2-Cys_6 binuclear cluster domain (Zn cluster). As shown in Fig. 2, Zn cluster, followed by C_2H_2 Zn fingers, are the prevalent DBDs in *T. melanosporum*, yet the total number of predicted TFs as well as their distribution among different classes do not differ significantly from those of other ascomycetes. The sole exceptions are the apparently higher number of TFs in *T. melanosporum* compared with *Schizosaccharomyces pombe*

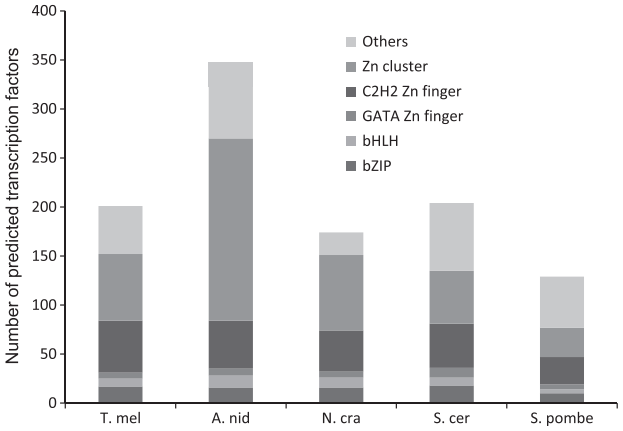

Fig. 2 Distribution of *Tuber melanosporum* (T. mel) transcription factors (TFs) among different structural categories (Stegmaier *et al.*, 2004) and comparison with the TF repertoire of unicellular and filamentous ascomycetes: *Saccharomyces cerevisiae* (S. cer) and *Schizosaccharomyces pombe* (S. pombe) (Beskow & Wright, 2006), *Aspergillus nidulans* (A. nid) (Wilson *et al.*, 2008) and *Neurospora crassa* (N. cra) (Borkovich *et al.*, 2004). The following types of DNA-binding domains (DBDs) were considered: Zn cluster; C_2H_2 Zn finger; GATA Zn finger; basic helix-loop-helix (bHLH); basic leucine zipper (bZIP); others, less represented types of DNA binding motifs ('Others').

and the nearly twofold lower number of TFs compared with *A. nidulans*, which displays a strong prevalence of Zn cluster transcription factors.

Functional assignment simply based on sequence similarity may lead to ambiguous results. This is especially true in the case of TFs, for which sequence similarity can be limited to (and biased by) highly conserved DBDs. We addressed this potential problem by performing a separate set of BLAST searches leaving out the DBDs. Based on these additional BLASTP results, putative *Tuber* TFs were subdivided into three classes: proteins apparently unique to *Tuber* with homology restricted to the DBD were classified as HP-DBDs (42 proteins); proteins for which sequence similarity is also significant outside of the DBD were classified as CH-TFs (57 proteins), or as pTFs if a functionally characterized ortholog could be identified (102 proteins). Predicted *Tuber* TFs belonging to each of the above classes, a total of 201 proteins, along with information on the particular type of DBD, best-hit sequence are listed in Table S1. All of the TF genes thus predicted gave above-background hybridization signals in oligoarray and/or RNAseq transcriptome analyses in at least one of the three life cycle stages that were investigated (FLM, ECM, FB); expression data for each candidate sequence are also reported in Table S1.

In addition to conserved basal transcriptional regulators (e.g. the RNA polymerase III transcription factors TFIIIA and TFIIIB; see Table S1), a specific transcriptional control function could be assigned to 91 of the predicted *T. melanosporum* proteins, which are similar to known TFs from other fungi, mainly yeast, *N. crassa* and *A. nidulans*. As shown in Table 1, putative *Tuber* TFs were assigned to five groups based on the different cellular functions of their potential target genes: cell wall modification, development, cell cycle, metabolism, and response to stress and stimuli.

All TFs, which, based on information available in other ascomycetes (both saprotrophic and phytopathogenic) are known to be involved in the regulation of genes encoding plant cell wall-degrading enzymes, are present in the *Tuber* genome (Table 1); for example, the carbon catabolite repressor CreA, the repressor Ace1, the pH regulator PacC, the Hap complex, the nitrogen regulator AreA, and the activator of genes encoding cellulose/hemicellulose degrading enzymes XlnR (Aro *et al.*, 2005) (Fig. 3a). One exception is the transcription factor Ace2, which at a variance with the situation in *Trichoderma reesei* and *N. crassa*, appears to be absent in *Tuber*. However, the lack of Ace2 was previously reported also in *A. nidulans* and an Ace2 homolog appears to be missing in *A. niger* and *Magnaporthe grisea* as well (Aro *et al.*, 2005), suggesting that different transcriptional regulators are likely to substitute for Ace2 function in these fungi. Of note, XlnR and the cutinase enzyme regulator Ctf1 are both preferentially expressed in mycorrhiza (ECM : FLM ratio *c.* 6 and *c.* 3, respectively; see Fig. 3a and Table S1).

Several homologs of *Neurospora* TF genes involved in development (e.g. hyphal growth and asexual/sexual development; Colot *et al.*, 2006) are present in the *Tuber* genome, along with regulators of FB development and pheromone response (e.g. SteA, the ortholog of the yeast regulator Ste12; StuA; NosA; Ste11; and the mating type protein Mat1-2-1 (Martin *et al.*, 2010) (see Table 1)). Curiously, Mat1-2-1 and SteA are both weakly expressed in FBs, while an ortholog of the *Aspergillus* NsdD gene (encoding a GATA-type transcription factor required for sexual development) could not be found in the *Tuber* genome. However, the homologs of other TFs that are essential for FB or conidiophore formation in *N. crassa* (e.g. the White Collar proteins WC-1 and WC-2) and *A. nidulans* (e.g. the general amino acid control activator CpcA, the G-protein modulator FlbA and the TFs FlbB, FlbC, and FlbD) are present in *T. melanosporum* (Poeggeler *et al.*, 2006).

Homologs of the main regulators of nitrogen and carbon metabolism are encoded by the *T. melanosporum* genome. These are accompanied by the homologs of general activators for the utilization of alternative nitrogen sources and pathway-specific activators for the metabolism of particular nitrogen sources. Of note, the activator of nitrate metabolism NirA is present in three copies in the *Tuber* genome. As shown in Table 1, regulators of aromatic amino acid biosynthesis were also identified, along with all known regulators of sulfur assimilation, which may be important for truffle flavor formation. In accordance with the sustained sulfur metabolism operating in fruiting bodies

Table 1 DNA-binding domain (DBD)-containing *Tuber melanosporum* transcription factors

Functional classes and gene names[a]	Putative gene product function[b]	Preferential expression[c] ECM	FB	FLM
Cell wall				
TmelXlnR	Activator of plant cell wall-degrading enzymes	↑		
TmelAce1	Repressor of plant cell wall-degrading enzymes		↓	
TmelCtf1	Transcriptional activator of cutinase genes	↑		
TmelMrgA	Regulation of cell wall integrity			
TmelRlmA	Maintenance of cell wall integrity			
Development				
Sexual development and fruiting body formation				
TmelStuA	Regulator of fruiting body and asexual development			
TmelSteA	Regulator of fruiting body development		↓	
TmelMat 1-2-1	MAT1-2-1 mating-type protein			
TmelMcm1	Regulator of pheromone response			
TmelSte11	Regulator of pheromone response			
TmelAsl1	Ascospore lethal 1			
TmelFf7	Female fertility	↑		
TmelSub2	Submerged protoperithecia			
TmelNosA	Number of sexual spores, regulator of sexual development			
TmelNdt80	Meiosis-specific transcription factor			
Asexual development and basal hyphal growth				
TmelLah3	Long aerial hyphae, regulator of asexual development		↓	
TmelTah1	Tall aerial hyphae, regulator of asexual development			
TmelCol27	Colonial, regulator of hyphal growth			
TmelCol21	Colonial, regulator of hyphal growth			
TmelVad5	Vegetative asexual development, regulator of hyphal growth and asexual development			
TmelVad3-1	Vegetative asexual development, regulator of hyphal growth and asexual development			↑
TmelVad3-2	Vegetative asexual development, regulator of hyphal growth and asexual development	↑		
TmelKal1	KALeidoscope-1, regulator of hyphal growth and asexual development			
TmelAbaA	Regulator of conidiation			↓
TmelDevR	Required for conidiophore development	↑		
TmelFlbB	Fluffy Low BrlA expression, regulator of asexual and conidiophore development			
TmelFlbC	Fluffy Low BrlA expression, regulator of asexual and conidiophore development			
TmelFlbD	Fluffy Low BrlA expression, regulator of asexual and conidiophore development	↑		
TmelOefC	Over expressed fluffy	↑		
TmelHms1	Regulator of cell morphology			
TmelCon7	Cell morphology regulator			
Other				
TmelAda1	All development altered, regulator of basal hyphal growth and asexual and sexual development			
TmelAda3	All development altered, regulator of basal hyphal growth and asexual and sexual development	↑		
Cell Cycle				
TmelSwi6	MBF complex, regulator of cell cycle			
TmelMbp1	MBF complex, regulator of cell cycle		↓	
TmelFkh2	Regulator of cell cycle			
TmelSep1	Regulator of cell cycle			
TmelSwi5	Regulator of cell cycle			
TmelYox1	Regulator of cell cycle			
TmelUme6	Key transcriptional regulator of early meiotic genes			
TmelRme1	Regulator of cell cycle: prevents meiosis, promotes mitosis	↑		
Metabolism				
Carbon				
TmelCreA	Major carbon catabolite repression protein			
TmelNrg1	Carbon catabolite repression			
TmelAzf1	Glucose responsive transcription factor		↓	
TmelRgt1	Glucose responsive transcription factor			
TmelFacB	Activates transcription of genes required for acetate utilization			
TmelAcuM	Positive regulator of gluconeogenesis			
TmelAcuK	Positive regulator of gluconeogenesis		↑	

Table 1 (Continued)

Functional classes and gene names[a]	Putative gene product function[b]	Preferential expression[c] ECM	FB	FLM
Nitrogen				
TmelAreA	Major, positively acting, nitrogen regulatory protein		↓	
TmelTamA	Positively acting nitrogen regulatory protein		↓	
TmelAreB	Major nitrogen catabolite repression protein		↓	
TmelMeaB	Nitrogen catabolite repression			
TmelGlcD	Regulation of nitrogen metabolism			
TmelCpcA	Cross pathway control protein, amino acid metabolism; regulator of fruiting body development			
TmelNirA-1	Pathway specific, positively acting nitrate regulatory protein			
TmelNirA-2	Pathway specific, positively acting nitrate regulatory protein			
TmelNirA-3	Pathway specific, positively acting nitrate regulatory protein	↑		
TmelAro80	Pathway specific, positively acting aromatic amino acid regulatory protein			
TmelUay1	Pathway specific, positively acting purine regulatory protein		↑	
TmelAmdX	Pathway specific, positively acting acetamidase regulatory protein			
TmelArcA	Pathway specific, positively acting arginine regulatory protein	↑		
TmelPpr1	Pathway specific, positively acting pyrimidine regulatory protein			
TmelLys14	Activator of lysine biosynthesis pathway			
Sulfur				
TmelMet32	Activator of sulfur metabolism		↓	
TmelMetR	Activator of sulfur metabolism			
TmelCbf1	Activator of sulfur metabolism; centromere binding protein		↑	
Lipids				
TmelOaf1	Activator of fatty acid metabolism			
TmelFarA	Activator of short- and long-chain fatty acid metabolism			↓
TmelFarB	Activator of short-chain fatty acid metabolism		↑	
TmelUpc2	Sterol regulatory element binding protein			
Other				
TmelPalcA	Transcription factor, involved in regulation of phosphate acquisition		↓	
TmelHapB	CCAAT binding complex, subunit B	↑		
TmelHapC	CCAAT binding complex, subunit C			
TmelHapE	CCAAT binding complex, subunit E			
TmelHapX	CCAAT binding complex, subunit X; iron-responsive factor			
Stress and stimuli response				
TmelSfl1	Activator of stress responsive genes			
TmelFcr3	FluConazole resistance protein 3		↓	
TmelSfp1	Cell cycle regulator in response to nutrients and stress			
TmelAsg1	Regulator of stress response and drug resistance			
TmelRpn4	Regulator of response to stress			↓
TmelHsf1	Heat shock transcription factor			
TmelAp1	Regulator of oxidative stress tolerance			
TmelSkn7	Response to osmotic and oxidative stress			
TmelPacC	Activator of alkaline-induced genes; repressor of acid-induced genes			
TmelHaa1	Weak acid stress response regulator			
TmelCrz1	Activator of genes involved in stress response			
TmelZap1	Activator of zinc responsive genes	↑		
TmelHacA	Unfolded protein response			
TmelWc1	Light response and circadian rhythm regulator			
TmelWc2	Light response and circadian rhythm regulator			
TmalSltA	Response to abiotic stress regulator			↑

[a]*T. melanosporum* DNA binding domain-containing transcription factors were grouped into major functional classes based on homology with functionally characterized TFs from other fungi; gene names were derived from those of the corresponding homologs (see Supporting Information Table S1 for further sequence information).
[b]Specific putative function of *Tuber* TFs as deduced from the known function of their characterized homologs.
[c]Arrows indicate expression ratios, derived from oligoarray and RNA-seq data, ≥ 2 (↑) or ≤ 0.5 (↓) in one of the investigated lifecycle stages with respect the other two (ECM, ECMs; FB: fruiting bodies; FLM: free-living mycelium); TAT-positive TFs are in bold.

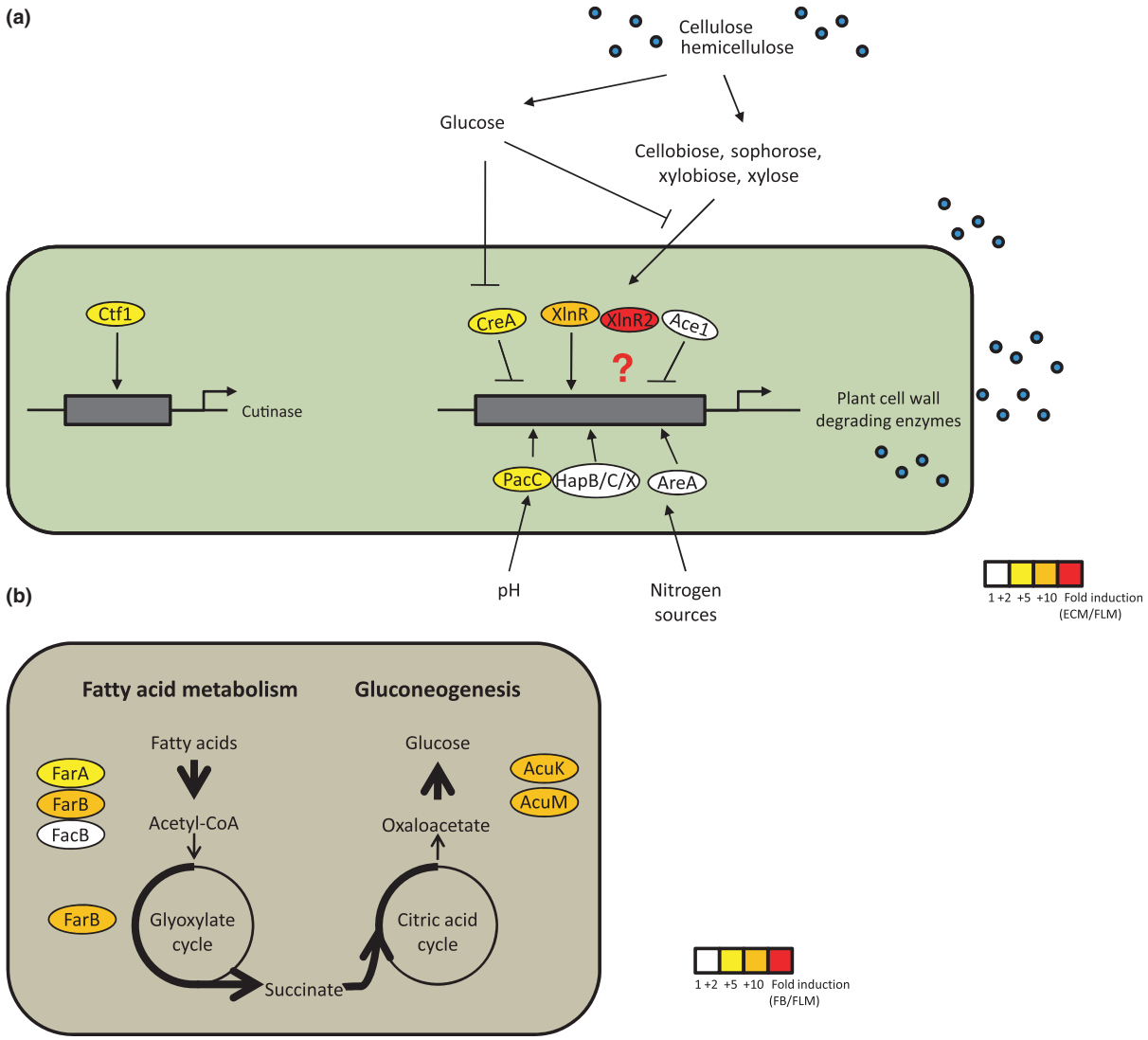

Fig. 3 Transcription factors (TFs) differentially expressed, and potentially involved in specific stages of the *Tuber* life cycle. (a) TFs regulating the expression of cell wall-degrading enzymes are preferentially expressed in ectomycorrhizas (ECMs). Also shown in this scheme is the *Tuber*-unique DBD-containing hypothetical protein (HP-DBD) encoded by GSTUMT00012451001, a gene that is strongly up-regulated in ECM and that was designated as XlnR2 because of its similarity to the cellulase/hemicellulase gene activator XlnR. (b) A subset of TFs preferentially expressed in fruiting bodies (FBs) positively regulate the expression of metabolic enzymes involved in fatty acid metabolism, including the glyoxylate cycle and gluconeogenesis, two pathways known to be particularly sustained in truffles (Abbà *et al.*, 2007). By comparison, only one regulator of sulfur metabolism (Cbf1) and no TF potentially involved in sexual development were found to be differentially up-regulated by more than twofold in FBs. TFs are represented with a false-color code (see the color scale inset) indicating their relative expression levels in ECMs or FBs compared with free-living mycelium (FLM, see text and Supporting Information, Table S1 for more information on these TFs).

(Martin *et al.*, 2010), the mRNAs for the latter components, which include the Cbf1 activator, exhibit a markedly preferential expression in fruiting bodies. The Met30 repressor is also present in the *Tuber* genome, but it is not listed in Table 1 because it does not bind to DNA directly.

Finally, regulators of 'stress and stimuli' responses, including the ECM-up-regulated activator of zinc homeostasis genes (Zap1), were identified. Genes coding for the major regulator of the unfolded protein response (HacA) are apparently amplified, with as many as 18 copies in the *Tuber* genome. However, careful inspection of the corresponding gene models showed that 17 putative HacA genes are severely truncated, not expressed, and/or very close to transposable elements, suggesting that all but one of them (TmelHacA) are putative pseudogenes (Table S1).

Functional screening and validation of truffle transcriptional activators in yeast

To validate functionally some of the TFs predicted earlier in the paper and to discover new transcriptional activators, bioinformatic analysis was coupled to the TAT assay, a

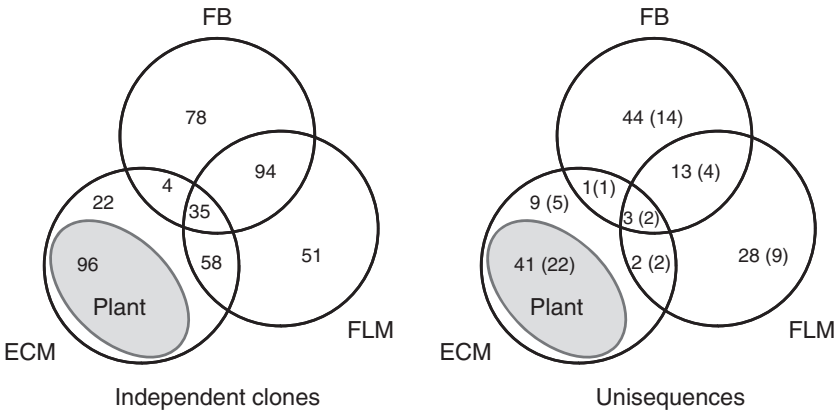

Fig. 4 Venn diagram representation of the number of independent clones (left) and corresponding unisequences (right) isolated from the transcriptional activator trap (TAT) screening of the free-living mycelium (FLM), fruiting body (FB) and ectomycorrhiza (ECM) libraries. The number of DNA-binding domain (DBD)-containing clones are shown between brackets. Sequences of plant origin retrieved from the screening of the ECM library are shown on a green background.

heterologous gene transactivation screen in yeast (Fig. 1). To this end, cDNA libraries prepared from the three life cycle stages of *T. melanosporum* were transferred to the pDEST32 yeast expression vector, and transformants bearing an in-frame fusion of the vector-provided yGAL4-DBD with a *Tuber* transcriptional activator were positively selected and monitored via reporter gene transactivation (see Fig. S1 for representative examples). At least 500 000 cDNA clones from each library were screened in this way and the inserts of positive selected clones were sequenced. A total of 438 sequences (160 from the FLM, 126 from the FB, and 152 from the ECM library) gave good-quality results and were retained for subsequent analysis. They were organized into 53 contigs and 88 singletons for a total of 141 unisequences (see Fig. 4). Most of these sequences (100) were of fungal origin, and 37 of them contained a recognizable DBD. Thirty-three of the latter sequences matched *in silico* predicted TFs, while the remaining four escaped detection because the DBD was missing as a result of misprediction of the first exon, or because only a partial DBD was present (see Table S1). As expected, a number of positive clones retrieved from the TAT screen of the ECM library (a total of 41 sequences) resembled plant polypeptides: 22 sequences bore a recognizable DBD and 12 of them shared a significant sequence similarity with plant transcriptional activators involved in plant–pathogen interaction (Fig. 5, Table S2). The remaining 82 TAT-positive sequences lacked a recognizable DBD and were collectively designated as 'unconventional activators' (see the paragraph 'Unconventional activators').

Among the functionally validated *Tuber* transcription factors homologous to TFs previously characterized in other fungi (pTFs; 22 sequences), we found two regulators of cell wall integrity (TmelAce1 and TmelMrgA), seven regulators of development and cell cycle genes, six stress response TFs, and seven regulators of metabolism (Table 1).

Five cDNAs, all belonging to the pTF class, were independently isolated > 20 times in the TAT assay and accounted for > 30% of the entire set of sequenced clones:

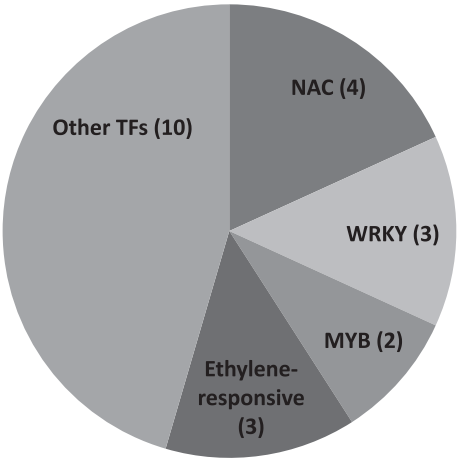

Fig. 5 DNA-binding domain (DBD)-containing transcription factors (TFs) of the plant (hazelnut) mycorrhizal partner retrieved from the transcriptional activator trap (TAT) screening of the ectomycorrhiza (ECM) library. TFs involved in microbial pathogen defense (Stracke *et al.*, 2001; Shi & Shi, 2004; Olsen *et al.*, 2005) are emphasized.

TmelCpcA (49 clones), TmelMetR (25 clones), TmelKal1 (22 clones), TmelMeaB (27 clones), plus the ethylene-responsive plant TF RAP2.12 (24 clones) (see Tables S1, S2). Analysis of the TmelMeaB contig revealed an expressed sequence tag (EST)/mRNA population which differed from that associated with the predicted gene model (GSTUMT00003085001) at the level of the 5′-splice site of the last intron. In fact, the initial gene model was predicted on the basis of the EST population produced by the downstream splice site (i.e. with a retained partial intron of 14 nt), while eight sequences mapping to this region, picked with the TAT screening, all contained the upstream splice site. Analysis of all the available ESTs, coupled with RNAseq data from FLM and FB, showed that one-third of the MeaB cDNAs had a retained partial intron, while two-thirds of them belonged to the same population identified with the TAT screening. Alternative splicing at these two sites thus produces two forms of MeaB with different

Table 2 DNA-binding domain (DBD)-lacking, unconventional activators retrieved from the transcriptional activator trap (TAT) screen[a]

Subcellular localization and description[b]	Gene model	*Saccharomyces cerevisiae* homolog
Nuclear proteins		
Mitogen-activated protein kinase	GSTUMT00005852001	Hog1
DNA-directed RNA polymerase II subunit RPB1	GSTUMT00009762001	Rpo21
Cryptic loci regulator 2	GSTUMT00004358001	
Mitogen-activated protein kinase MAF1	GSTUMT00005390001	Maf1
Casein kinase II beta 2 subunit	GSTUMT00011807001	
Serine/threonine-protein kinase KSP1	GSTUMT00006520001	Ksp1
DASH complex subunit ask1	GSTUMT00005863001	Ask1
Ubiquitin thiolesterase (OtuB1)	GSTUMT00000718001	
Mixed localization proteins		
E3 ubiquitin-protein ligase NEDD4	GSTUMT00011744001	Rsp5
Actin	GSTUMT00006162001	Act1
EH Domain and endocytosis-related protein	GSTUMT00009307001	Ede1
Alpha-actinin	GSTUMT00006491001	
Charged multivesicular body protein 5	GSTUMT00012039001	Vps60
Calmodulin	GSTUMT00010275001	Cmd1
Low-temperature viability protein ltv1	GSTUMT00001950001	Ltv1
Peptide methionine sulfoxide reductase msrB/msrA[c]	GSTUMT00004791001	Mxr1
Phospholipase D1 (PLD1), putative	GSTUMT00003661001	Spo14
Mitosis inhibitor protein kinase SWE1	GSTUMT00005490001	Swe1
vts1 protein	GSTUMT00007890001	Vts1
Ankyrin repeat protein	GSTUMT00003500001	Nas6

[a]TAT-positive unconventional activators retrieved from functional screening of the *Tuber melanosporum* free-living mycelium (FLM), fruiting body (FB) and ectomycorrhiza (ECM) libraries in yeast (see Supporting Information, Table S3 for further information).
[b]The subcellular (nuclear or mixed) localization of homologous proteins from other organisms along with the gene name of the yeast homolog, when available, are indicated.
[c]Unconventional activator differentially expressed in FB, with FB : FLM and FB : ECM expression ratios of 40.0-2.1 and 747.7-5.7 (oligoarrays-RNAseq), respectively.

C-terminal sequences. The form encoded by the mRNA lacking the retained partial intron is 61 amino acids longer, is more similar to MeaB fungal proteins, and is thus probably the only one that is transcriptionally active.

Fifteen of the TAT-positive clones were either *Tuber*-unique HP-DBDs (six) or CH-TFs (nine). Since they all met the basic criteria of our analysis (i.e. the presence of a DBD and the ability to activate reporter gene expression in yeast, along with above-background expression levels in at least one stage of the *Tuber* life cycle), we consider the corresponding proteins as newly identified, *Tuber*-unique (HP-DBDs) or general filamentous ascomycete (CH-TFs) TFs (see Table S1).

The TAT screen thus allowed the functional validation of about one-fifth of *in silico* predicted TFs, the adjustment of three gene model(s), and the *de novo* identification of four DBD-containing transcriptional activators. This was accompanied by the identification of at least 22 *bona fide* plant TFs potentially bearing on the response of the host plant to *T. melanosporum* infection and by the discovery of 82 unconventional putative transcriptional activators (see the paragraph 'Unconventional activators').

Unconventional activators

DNA-binding domain-lacking sequences similar to previously described nuclear or mixed localization proteins as well as to proteins apparently functioning outside of the nucleus were also found among the TAT-positive clones. These included 20 proteins, eight nuclear and 12 with a mixed (nuclear/cytoplasmic) localization, called 'unconventional activators' (UAs), for which a direct (or indirect) role in transcriptional regulation has previously been demonstrated (Table 2), and 43 proteins, called 'putative unconventional activators' (PUAs) with no previously documented involvement in transcription (Table S3).

Among the UAs we found various kinases, ubiquitinating enzymes, endocytotic and actin-related proteins, some of which are known to shuttle between the cytoplasm and the nucleus, where they act as scaffolds for the recruitment of the basal transcription apparatus or play a role in chromatin modification (Pilecka *et al.*, 2007; Zheng *et al.*, 2009). Various nuclear actions have been reported for these proteins, which are often found associated with actively transcribed mammalian genes. These include transcription stimulation,

chromatin remodeling, RNA/protein export from the nucleus in the case of actin, and enhanced transactivation by the nuclear receptor activator SRC-2 in the case of alpha-actinin (Zheng *et al.*, 2009). Two endocytotic proteins with a previous record of nuclear actions were also found among the TAT-positive clones: TmelEde1, which is homologous to human Eps15p (Vecchi *et al.*, 2001), and TmelVps60, the *Tuber* homolog of charged multivesicular body protein 5 (CHMP5), which has been shown to interact with chromatin-remodeling enzymes as well with proteins involved in nuclear sumoylation and with the RNA polymerase III transcription factor IIIA in human cells (Tsang *et al.*, 2006).

Perhaps more problematic are the 'putative unconventional activators'. In fact, only half of these proteins – six of which are conserved hypothetical proteins and two of which are *Tuber*-unique proteins – bear a recognizable NLS and can thus be hypothesized to be capable of localizing to the nucleus autonomously. Indirect nuclear localization mechanisms (e.g. 'piggy back' translocation) along with some false positives (i.e. cytoplasmic proteins that function as transcriptional activators when artificially forced into the nucleus by the vector-borne NLS) can instead be hypothesized for the other 35 TAT-positive polypeptides. As previously noted, however, there is an increasing number of apparently cytoplasmic proteins for which a genuine nuclear function has been documented (Shi & Shi, 2004; Pilecka *et al.*, 2007). In keeping with recent findings in yeast (Hall *et al.*, 2004), among *Tuber* PUAs we found metabolic enzymes, but also proteins involved in vesicular transport. Plant-derived UAs (e.g. the nuclear signaling decoder calmodulin, involved in the response to pathogen invasion) and PUAs were also retrieved from TAT analysis of the ECM library and are listed in Table S2. A detailed analysis of the ability of TAT-positive PUAs to enter the nucleus autonomously is under way.

Transcription factors preferentially expressed in ECMs and FBs

Preferential expression of all predicted TFs in the ectomycorrhizal and/or FB stage of the *Tuber* life cycle was investigated by microarray hybridization and RNA sequencing (see the 'Materials and Methods' section and (Martin *et al.*, 2010) for details). Twenty-five TFs were found to be up-regulated more than fivefold in ECM (16) or FB (nine) relative to FLM, with up-regulation levels as high as 49-fold and 4500-fold for the ECM and the FB stage, respectively (Table 3). A homolog of the stress response regulator Rpn4, two conserved hypothetical TFs and a *Tuber*-specific Zn cluster TF were up-regulated more than fivefold in both life cycle stages. Five of the above TFs were independently retrieved from both *in silico* analysis and TAT screening, while seven of them were individually subjected, and scored positive to the TAT assay following prioritarization by expression profiling.

Transcription factors with no similarity to functionally characterized proteins were prevalent among transcriptional activators specifically up-regulated in ECMs. Thus, only indirect inferences regarding the function and potential targets of these TFs can be made at present. For example, the hypothetical Zn cluster DBD-containing protein encoded by GSTUMT00012451001 – one of the top up-regulated TFs in ECM – is distantly related to XlnR, an activator of plant cell wall-degrading enzymes (cellulases/hemicellulases) in other fungi. An ortholog of the XlnR activator, named TmelXlnR, is present in the *Tuber* genome and is up-regulated by more than fivefold in ECM compared with FLM. However, its expression levels are 10-fold lower than those for the transcript encoded by GSTUMT00012451001, a novel putative activator that we designated as XlnR2 (Fig. 3a, Table S1), whose polysaccharide substrates and physiological role in *Tuber* are presently unknown.

Among ECM-preferential TFs with a functionally characterized homolog in other fungi, we found the regulator of cell morphology TmelHms1, the regulators of asexual development TmelFlbD and TmelDevR, and the regulator of sexual development TmelFf7. Interestingly, TmelFf7 as well as its *Neurospora* homolog contain the maltose acetyltransferase domain typical of the *Rhizobium* nodulation factor NodL. The stress response regulators *TmelSFP1* and *TmelSFL1* were also found to be overexpressed in ECM, suggesting that control of stress responsiveness and morphological differentiation may be jointly regulated processes during ECM formation. Of note, the homologs of TmelSfp1 and of the conserved C_2H_2-Zn finger TF GSTUMT00007914001 are also up-regulated in ECMs from the basidiomycete *Laccaria bicolor*.

Nine predicted TFs, six containing a Zn cluster DBD, are preferentially expressed in FBs. Five of these TFs have a functionally characterized homolog in other fungi: TmelAbaA, a regulator of asexual development; the positive regulators of gluconeogenesis TmelAcuM and TmelAcuK; the sulfur metabolite activator TmelCbf1; and the activator of short-chain fatty acid metabolism TmelFarB. FB-preferential expression of the latter four regulators fits with the notion that gluconeogenesis, especially via fatty acid utilization through the glyoxylate cycle (Fig. 3b), and sulfur assimilation are highly sustained metabolic pathways in *Tuber* FBs. The remaining four FB-preferential activators are either 'conserved hypothetical TFs' (two) or *Tuber*-unique HP-DBDs (two). One of the latter proteins, bearing a Zn cluster DBD, is encoded by the second most up-regulated gene (GSTUMT00009814001) in FBs.

Discussion

This work describes the repertoire of TFs, the so-called 'regulome', of a plant-symbiotic ascomycete producing highly prized subterranean FBs. A total of 201 sequences

Table 3 Transcription factors (TFs) differentially up-regulated in ectomycorrhizas (ECMs) or fruiting bodies (FBs)[a]

		RNAseq (expression ratios)		Oligoarray (expression ratios)		
Gene model	Name	ECM : FLM	FB : FLM	ECM : FLM	FB : FLM	Putative function
GSTUMT00004347001	*TmelXLNR*	**5.4**	0.1	**6.2**	0.0	*Cell wall*
GSTUMT00005078001	*TmelABAA*	nd	nd	**15.9**	> **15.1**	*Development and*
GSTUMT00007803001	*TmelDEVR*	**13.4**	0.4	3.5	1.3	*hyphal growth*
GSTUMT00008124001	*TmelFF7*	**9.2**	0.5	**5.4**	1.2	
GSTUMT00005459001	*TmelHMS1*	**7.5**	1.6	1.8	1.4	
GSTUMT00009849001	*TmelFLBD*	**6.8**	0.0	**7.3**	0.0	
GSTUMT00008596001	*TmelACUM*	1.2	**13.1**	1.0	1.4	*Carbon metabolism*
GSTUMT00012254001	*TmelFARB*	0.7	**12.5**	0.4	2.2	
GSTUMT00000111001	*TmelACUK*	0.7	**6.5**	0.4	**11.7**	
GSTUMT00000814001	*TmelCBF1*	3.2	**14.0**	2.0	**5.2**	*Sulfur metabolism*
GSTUMT00012172001	*TmelMET32*	**8.0**	0.0	1.7	0.0	
GSTUMT00006696001	*TmelRPN4*	**16.2**	**10.8**	3.2	1.6	*Stress response*
GSTUMT00004466001	*TmelSFL1*	**5.8**	0.6	1.4	0.8	
GSTUMT00008521001	*TmelSFP1*	**5.1**	2.0	1.2	1.8	
GSTUMT00007914001	CH-TF	**6.2**	3.6	**44.4**	**10.4**	*Conserved hypothetical*
GSTUMT00004451001	CH-TF	**48.8**	0.2	**33.7**	0.6	*transcription factors*
GSTUMT00006533001	CH-TF	**18.2**	**8.0**	**25.6**	**29.6**	
GSTUMT00009188001	CH-TF	1.3	**37.0**	0.7	**14.3**	
GSTUMT00011438001	CH-TF	3.9	**5.0**	1.9	2.3	
GSTUMT00006234001	CH-TF	**5.0**	3.9	**5.4**	3.3	
GSTUMT00000561001	CH-TF	**5.5**	0.3	**7.4**	1.4	
GSTUMT00004341001	CH-TF	**7.9**	3.1	4.5	3.7	
GSTUMT00009814001	HP-DBD	1.9	**180.3**	1.3	**> 4478.2**	*DNA-binding domain*
GSTUMT00004776001	HP-DBD	nd	nd	**8.4**	**84.5**	*(DBD)-containing*
GSTUMT00005836001	HP-DBD	**12.8**	**8.4**	**5.5**	2.1	*hypothetical proteins*
GSTUMT00011475001	HP-DBD	**8.2**	0.3	**7.4**	1.0	
GSTUMT00012451001	HP-DBD	**28.8**	0.0	**7.1**	0.1	
GSTUMT00006134001	HP-DBD	**5.2**	0.3	2.3	0.3	
GSTUMT00000279001	HP-DBD	**9.1**	1.6	4.1	1.7	

CH-TFs, conserved hypothetical transcription factors; HP-DBDs, DBD-containing hypothetical proteins; FLM, free-living mycelium.
[a]Transcription factors, CH-TFs and HP-DBDs, which, based on RNAseq or oligoarray data (or both), are up-regulated by at least fivefold in ECMs and/or FBs compared with FLM are reported. Instances of oligoarray ratios ≥ 5 (P-value < 0.05) with RNAseq ratio > 1, or oligoarray ratio > 1 with RNAseq ratio > 5, were both considered as positive hits. RNAseq expression ratios > 5 are in bold; oligoarray expression ratios > 5 with a significant (< 0.05) or nonsignificant (> 0.05) P-value are shown in black-bold and light gray-bold, respectively. Cases in which no RNAseq data were available for the reference FLM stage are indicated as 'nd'. TFs were classified according their putative function based on similarity with characterized fungal TFs; gene models corresponding to transcriptional activator trap (TAT)-validated proteins are in bold (see Table S1 for further details).

coding for putative and/or functionally validated TFs has been identified by the combined use of *in silico* prediction and a TAT assay conducted in the model ascomycete *S. cerevisiae*. The TAT assay allowed the validation of functionally about one-fifth of the *in silico* predicted *Tuber* TFs; a hit rate that is only twofold lower than that reported for the homologous validation of yeast TFs, where the lack of reporter gene activation for 83 out of 138 predicted yeast transcriptional activators was ascribed to the requirement for specific promoter elements, cofactors and/or growth conditions (Titz *et al.*, 2006). On the other hand, the totally unbiased nature of the TAT assay led to the *de novo* identification of four TFs along with a number of 'unconventional activators' that would have been impossible to predict simply based on sequence similarity. The same screening also allowed the assignment of a putative 'transcriptional activation function' to 12 hypothetical, functionally uncharacterized proteins (CH-TFs) with a widespread occurrence in fungi. Additionally, we verified that all the candidate TFs we identified are expressed in at least one stage of the *Tuber* life cycle, and 14% of them were found to be transcriptionally up-regulated in a stage-dependent manner in ECMs or FBs.

Most TFs are encoded by single-copy genes in the *T. melanosporum* genome. The sole exceptions (leaving out a few instances of TF pseudogenes) are the homolog of the *N. crassa* regulator of hyphal growth and asexual development *VAD3*, which is present in two copies, and the nitrate metabolism activator *NIRA*, which is present in three copies. The occurrence of extra copies of *NIRA* is potentially interesting in view of the fungal nitrogen–plant carbon trade taking place in ECMs and because nitrogen metabolism is likely to be altered during symbiosis establishment

and maintenance. In fact, the expression of nitrate assimilation enzymes has previously been shown to be modified in *Tuber borchii* ECMs and, unlike the situation in most fungi, nitrate assimilation enzymes appear not to be subjected to nitrogen catabolite repression in *Tuber* (Montanini *et al.*, 2006b; Guescini *et al.*, 2007).

The fact that > 50% of the *Tuber* TFs are homologous to previously characterized transcriptional regulators from other fungi served as a 'positive control' for our analysis and provided important clues regarding the gene targets and potential function of these TFs. This is especially true for some TFs preferentially expressed in ECMs (e.g. TmelXlnR, TmelCtf1, TmelAce1), whose homologs in other filamentous ascomycetes have been implicated in biochemical processes related to plant cell wall degradation and to the establishment of biotrophic (plant–pathogen) interactions (Fig. 3a); for example, genes with a role in the deconstruction of plant cell walls, host cell adhesion, detoxification of plant defense metabolites, and nutrient exchange, many of which are differentially up-regulated in *T. melanosporum* ECMs (Martin *et al.*, 2010). Similar considerations hold for a subset of TFs preferentially expressed in FBs (e.g. TmelACUM, TmelFARB, TmelACUK and TmelCBF1), which pinpointed potential target genes involved in metabolic pathways (e.g. gluconeogenesis, glyoxylate cycle and sulfur assimilation) that are particularly active in FBs (Fig. 3b) (Lacourt *et al.*, 2002; Gabella *et al.*, 2005; Abbà *et al.*, 2007; Martin *et al.*, 2010).

Identification of putative homologs of TFs acting on specific genes and pathways in other fungi may also prompt investigations aimed at verifying the occurrence of functionally related target genes in *Tuber*. This is the case, for example, of the two *Tuber* homologs of XlnR, a major regulator of sugar hydrolytic enzyme genes (endoglucanases, cellobiohydrolases and xylanases) in saprotrophic filamentous ascomycetes (Aro *et al.*, 2005). Three beta-1,4-endoglucanases are encoded by the *Tuber* genome and one of them (TmelCMC3) is strongly up-regulated (> 6000-fold) in mycorrhiza, similar to its *L. bicolor* ortholog (Martin *et al.*, 2008). In contrast, no cellobiohydrolase and one, but ECM-down-regulated, xylanase (endo-1,4-beta-xylanase precursor; TmelXylA) are present in *T. melanosporum*.

Similar considerations hold for the *Tuber* homolog of the cutinase enzyme regulator (Ctf1), for which no apparent cutinase gene could be identified in the *T. melanosporum* genome. It should be noted, however, that the physiological targets of mycorrhizal fungi are host plant roots, that do not contain cutin, and that the Ctf transcription factors initially characterized in *Nectria haematococca* are the orthologs of the lipid metabolism regulators FarA and FarB, with which they share a common DNA binding site. Indeed, comparative genomic analyses in *Aspergillus* spp. highlighted the presence of conserved FarA and FarB DNA binding sites upstream of cutinase, but also of lipid metabolism and peroxisomal enzyme genes (Hynes *et al.*, 2006). Interestingly, one of the genes most strongly up-regulated in *T. melanosporum* mycorrhiza codes for a secreted lipase/esterase (Martin *et al.*, 2010), which might be involved in host membrane remodeling.

Of note, about one-quarter of the FB- or ECM-up-regulated TFs (see Table 3) are among the top 10% most highly expressed genes in FBs or ECMs. This is a rather peculiar situation, considering that TF-encoding genes are usually expressed at very low levels. This might point to the high degree of functional specialization associated with different stages of the *Tuber* life cycle and warrants future investigations aimed at discovering the actual target genes and validating the biological processes associated with these TFs. A good case in point is represented by a few TFs whose homologs in other ascomycetes are implicated in asexual development and somatic spore (e.g. conidia) formation, a process whose existence in *Tuber* spp. is still very controversial (Urban *et al.*, 2004). Similar in-depth investigations, including the experimental identification of the DNA binding sites, will specifically address *Tuber*-unique TFs, such as the TAT-validated activator encoded by GSTUMT00009814001 – the second most expressed gene in FBs – as well as newly identified UAs. The latter components, especially those without a prior record of nuclear localization and activity (PUAs), represent another significant outcome of this analysis. Although a relatively high rate of false positives may be expected as a result of the presence of a vector-borne NLS, the occurrence of moonlighting proteins capable of functioning both in the cytoplasm and in the nucleus is in line with recent findings in yeast (and other organisms) as to the gene transactivation ability of otherwise cytoplasmic proteins, especially metabolic enzymes (Hall *et al.*, 2004; Pilecka *et al.*, 2007; Hu *et al.*, 2009). This points to the existence of an as yet largely unexplored set of hard-to-predict transcriptional regulators and raises the exciting possibility of a direct link between metabolism and transcription.

A final note regards the plant TFs retrieved from the TAT screen of the ECM library, 12 of which belong to specific classes of transcriptional regulators (e.g. NAC, WRKY, MYB and ER) that in plants attacked by phytopathogenic fungi have been shown to control various kinds of defense-related processes (Stracke *et al.*, 2001; Singh *et al.*, 2002; Olsen *et al.*, 2005). Another case in point is the presence among plant UAs of the hazelnut homolog of *Arabidopsis* calmodulin, a multifunctional protein that, despite its predominant cytoplasmic localization, is capable of entering the nucleus, where it has recently been shown to negatively regulate defense responses to pathogen invasion in a Ca^{2+}-dependent manner (Kim *et al.*, 2009). Plant TFs functionally identified in *T. melanosporum*/hazelnut mycorrhiza may thus open a new window on the mutualistic interaction between symbiotic fungi and their plant partners, and

provide important clues as to its mechanistic peculiarities compared with better characterized host and nonhost plant–pathogen interactions.

Acknowledgements

We thank Patrick Wincker and all the Génoscope team (Evry, France) for their invaluable contribution to the sequencing and assembly of the *T. melanosporum* genome. We are grateful to Gérard Chevalier (INRA-Universite' Blaise Pascal, Clermont-Theix, France) for *T. melanosporum* mycelia and fruiting bodies; Raffaella Balestrini (Istituto per la Protezione delle Piante, CNR, Torino, Italy) for mycorrhizal root tips; and Marian Walhout (University of Massachusetts Medical School, Worcester, USA) for yeast strain MaV103. Financial support from the Fondazione Cariparma (Parma, Italy) and from the Ministry of Education, University and Research (Rome, Italy; MIUR-PRIN 2007 program) to S.O. is gratefully acknowledged. B.M. was funded by a postdoctoral fellowship from the Fondazione Cariparma; E.L. is a graduate student in 'Biochemistry and Molecular Biology' at the University of Parma. Work in F.M.'s laboratory was supported by grants from the European Network of Excellence EVOLTREE and the Région Lorraine Council (Project FORBOIS). E.T. was funded by a joint PhD scholarship from the European project ENERGYPOPLAR and Région Lorraine.

References

Abbà S, Balestrini R, Benedetto A, Rottensteiner H, De Lucas JR, Bonfante P. 2007. The role of the glyoxylate cycle in the symbiotic fungus *Tuber borchii*: expression analysis and subcellular localization. *Current Genetics* **52**: 159–170.

Aro N, Pakula T, Penttila M. 2005. Transcriptional regulation of plant cell wall degradation by filamentous fungi. *FEMS (Federation of European Microbiological Societies) Microbiology Reviews* **29**: 719–739.

Balestrini R, Mainieri D, Soragni E, Garnero L, Rollino S, Viotti A, Ottonello S, Bonfante P. 2000. Differential expression of chitin synthase III and IV mRNAs in ascomata of *Tuber borchii* Vittad. *Fungal Genetics and Biology* **31**: 219–232.

Beskow A, Wright AP. 2006. Comparative analysis of regulatory transcription factors in *Schizosaccharomyces pombe* and budding yeasts. *Yeast* **23**: 929–935.

Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, Seiler S, Bell-Pedersen D, Paietta J, Plesofsky N *et al.* 2004. Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. *Microbiology and Molecular Biology Reviews* **68**: 1–108.

Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. 2006. A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proceedings of the National Academy of Sciences, USA* **103**: 10352–10357.

Gabella S, Abba S, Duplessis S, Montanini B, Martin F, Bonfante P. 2005. Transcript profiling reveals novel marker genes involved in fruiting body formation in *Tuber borchii*. *Eukaryotic Cell* **4**: 1599–1602.

Guescini M, Zeppa S, Pierleoni R, Sisti D, Stocchi L, Stocchi V. 2007. The expression profile of the *Tuber borchii* nitrite reductase suggests its positive contribution to host plant nitrogen nutrition. *Current Genetics* **51**: 31–41.

Hall DA, Zhu H, Zhu X, Royce T, Gerstein M, Snyder M. 2004. Regulation of gene expression by a metabolic enzyme. *Science* **306**: 482–484.

Horton P, Nakai K. 1997. Better prediction of protein cellular localization sites with the k nearest neighbors classifier. *Proceedings/ International Conference on Intelligent Systems for Molecular Biology* **5**: 147–152.

Hu S, Xie Z, Onishi A, Yu X, Jiang L, Lin J, Rho HS, Woodard C, Wang H, Jeong JS *et al.* 2009. Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. *Cell* **139**: 610–622.

Hynes MJ, Murray SL, Duncan A, Khew GS, Davis MA. 2006. Regulatory genes controlling fatty acid catabolism and peroxisomal functions in the filamentous fungus *Aspergillus nidulans*. *Eukaryotic Cell* **5**: 794–805.

Kim MC, Chung WS, Yun DJ, Cho MJ. 2009. Calcium and calmodulin-mediated regulation of gene expression in plants. *Molecular Plant* **2**: 13–21.

Lacourt I, Duplessis S, Abba S, Bonfante P, Martin F. 2002. Isolation and characterization of differentially expressed genes in the mycelium and fruit body of *Tuber borchii*. *Applied and Environmental Microbiology* **68**: 4574–4582.

Martin F, Aerts A, Ahrén D, Brun A, Danchin EGJ, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V *et al.* 2008. The genome sequence of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* **452**: 88–93.

Martin F, Kohler A, Duplessis S. 2007. Living in harmony in the wood underground: ectomycorrhizal genomics. *Current Opinion in Plant Biology* **10**: 204–210.

Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, Montanini B, Morin E, Noel B, Percudani R *et al.* 2010. Perigord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* **464**: 1033–1038.

Masoudi-Nejad A, Tonomura K, Kawashima S, Moriya Y, Suzuki M, Itoh M, Kanehisa M, Endo T, Goto S. 2006. EGassembler: online bioinformatics service for large-scale processing, clustering and assembling ESTs and genomic DNA fragments. *Nucleic Acids Research* **34**: W459–W462.

Montanini B, Gabella S, Abba S, Peter M, Kohler A, Bonfante P, Chalot M, Martin F, Ottonello S. 2006a. Gene expression profiling of the nitrogen starvation stress response in the mycorrhizal ascomycete *Tuber borchii*. *Fungal Genetics and Biology* **43**: 630–641.

Montanini B, Viscomi AR, Bolchi A, Martin Y, Siverio JM, Balestrini R, Bonfante P, Ottonello S. 2006b. Functional properties and differential mode of regulation of the nitrate transporter from a plant symbiotic ascomycete. *Biochemical Journal* **394**: 125–134.

Olsen AN, Ernst HA, Leggio LL, Skriver K. 2005. NAC transcription factors: structurally distinct, functionally diverse. *Trends in Plant Science* **10**: 79–87.

Pilecka I, Banach-Orlowska M, Miaczynska M. 2007. Nuclear functions of endocytic proteins. *European Journal of Cell Biology* **86**: 533–547.

Poeggeler S, Nowrousian M, Kueck U. 2006. Fruiting-body development in ascomycetes. In: Kues U, Fischer R, (eds). *The Mycota*, Vol. I. Berlin, Heidelberg, Germany: Springer-Verlag, 325–355.

Shi Y, Shi Y. 2004. Metabolic enzymes and coenzymes in transcription–a direct link between metabolism and transcription? *Trends in Genetics* **20**: 445–452.

Singh K, Foley RC, Onate-Sanchez L. 2002. Transcription factors in plant defense and stress responses. *Current Opinion in Plant Biology* **5**: 430–436.

Stegmaier P, Kel AE, Wingender E. 2004. Systematic DNA-binding domain classification of transcription factors. *Genome Informatics* **15**: 276–286.
Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* **4**: 447–456.
Titz B, Thomas S, Rajagopala SV, Chiba T, Ito T, Uetz P. 2006. Transcriptional activators in yeast. *Nucleic Acids Research* **34**: 955–967.
Tsang HT, Connell JW, Brown SE, Thompson A, Reid E, Sanderson CM. 2006. A systematic analysis of human CHMP protein interactions: additional MIT domain-containing proteins bind to multiple components of the human ESCRT III complex. *Genomics* **88**: 333–346.
Urban A, Neuner-Plattner I, Krisai-Greilhuber I, Haselwandter K. 2004. Molecular studies on terricolous microfungi reveal novel anamorphs of two *Tuber* species. *Mycological Research* **108**: 749–758.
Vecchi M, Polo S, Poupon V, van de Loo JW, Benmerah A, Di Fiore PP. 2001. Nucleocytoplasmic shuttling of endocytic proteins. *Journal of Cell Biology* **153**: 1511–1517.
Walhout AJ, Vidal M. 2001. High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. *Methods* **24**: 297–306.
Wilson D, Charoensawan V, Kummerfeld SK, Teichmann SA. 2008. DBD–taxonomically broad transcription factor predictions: new content and functionality. *Nucleic Acids Research* **36**: D88–D92.
Ye R, Yao QH, Xu ZH, Xue HW. 2004. Development of an efficient method for the isolation of factors involved in gene transcription during rice embryo development. *Plant Journal* **38**: 348–357.
Zaretsky M, Sitrit Y, Mills D, Roth-Bejerano N, Kagan-Zur V. 2006. Differential expression of fungal genes at preinfection and mycorrhiza establishment between *Terfezia boudieri* isolates and *Cistus incanus* hairy root clones. *New Phytologist* **171**: 837–845.
Zheng B, Han M, Bernier M, Wen JK. 2009. Nuclear actin and actin-binding proteins in the regulation of transcription and gene expression. *FEBS Journal* **276**: 2669–2685.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Examples of reporter gene activation results produced by the transcriptional activator trap screening of *Tuber melanosporum* transcription factors in yeast.

Table S1 List of *Tuber* transcription factors with general information on each predicted protein

Table S2 List of plant transcription factors, including 'unconventional activators' (UAs) and 'putative unconventional activators' (PUAs) retrieved from the transcriptional activator trap screening of the ectomycorrhiza library

Table S3 Transcriptional activator trap-positive, *Tuber melanosporum* 'unconventional activators' (UAs) and 'putative unconventional activators' (PUAs)

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.