A eukaryotic nicotinate-inducible gene cluster: convergent evolution in fungi and bacteria

Judit Ámon1,†, Rafael Fernández-Martín2,1,†, Eszter Bokor1, Antonietta Cultrone2,6, Joan M. Kelly3,4, Michel Flipphi2,‖, Claudio Scazzocchio2,3,4,5 and Zsuzsanna Hamari1,2

1Department of Microbiology, University of Szeged Faculty of Science and Informatics, Szeged, Hungary (present address of ZH)
2Institute de Génétique et Microbiologie, Université Paris-Sud, Orsay, France
3Department of Biology, University of Essex, Colchester, UK
4Department of Microbiology, Imperial College, London, UK (present address of CS)
5Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France (present address of CS)
6Department of Genetics and Evolution, University of Adelaide, Adelaide, Australia
7Institute of Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France (present address of CS)

Nicotinate degradation has hitherto been elucidated only in bacteria. In the ascomycete Aspergillus nidulans, six loci, hxnS/AN9178 encoding the molybdenum cofactor-containing nicotinate hydroxylase, AN11197 encoding a Cys2/His2 zinc finger regulator HxnR, together with AN11196/hxnZ, AN11188/hxnY, AN11189/hxnP and AN9177/hxnT, are clustered and stringently co-induced by a nicotinate derivative and subject to nitrogen metabolite repression mediated by the GATA factor AreA. These genes are strictly co-regulated by HxnR. Within the hxnR gene, constitutive mutations map in two discrete regions. Aspergillus nidulans is capable of using nicotinate and its oxidation products 6-hydroxynicotinic acid and 2,5-dihydroxypyridine as sole nitrogen sources in an HxnR-dependent way. HxnS is highly similar to HxA, the canonical xanthine dehydrogenase (XDH), and has originated by gene duplication, preceding the origin of the Pezizomycotina. This cluster is conserved with some variations throughout the Aspergillaceae. Our results imply that a fungal pathway has arisen independently from bacterial ones. Significantly, the neo-functionalization of XDH into nicotinate hydroxylase has occurred independently from analogous events in bacteria. This work describes for the first time a gene cluster involved in nicotinate catabolism in a eukaryote and has relevance for the formation and evolution of co-regulated primary metabolic gene clusters and the microbial degradation of N-heterocyclic compounds.

1. Introduction

Filamentous ascomycetes comprise metabolically versatile saprophytes that can use a large variety of metabolites as nitrogen and/or carbon sources. The utilization of nicotinic acid has been studied in bacteria, but it has only been addressed in a eukaryotic microorganism by our early work in Aspergillus nidulans. An enzyme of the xanthine dehydrogenase (XDH) group [1–3] is necessary for this process. Strains mutant in the cnx (cnxABC, cnxE, cnxF, cnxG and cnxH) or hxB genes cannot use nicotinate. The cnx genes are required for the synthesis of the molybdenum cofactor (MOCO) common to XDH and nitrate reductase [4,5]. The HxB protein catalyses the sulfuration of the Mo(VI), essential for the activity of the enzymes of the XDH group [5,6].

Two enzymes of the XDH family have been described in A. nidulans. Purine hydroxylase I (PHI, HxA encoded by the hxA gene) is a typical XDH [7–9]. Purine hydroxylase II (PHII, HxnS; see below) has unprecedented substrate
specificity. Hypoxanthine, but not xanthine, serves as a substrate of PHII. It accepts nicotinate as a substrate and catalyses the first step of nicotinate catabolism [1,7,10]. Table 1 presents some kinetic parameters for PHI (HxA) and PHII (HxnS) summarized from the relevant literature.

PHII is absent in mycelia grown on nitrogen sources generally considered non-repressive. It is apparently induced by nicotinate but it is also present in nitrogen-starved mycelia [1]. The physiological inducer is either 6-OH nicotinate and/or a metabolite further along the nicotinate utilization pathway [12]. The expression of PHII is not under the control of UaY, the transcription factor specific for the expression of the genes in the purine utilization pathway including \( hxA \) [13–15].

Concentrations of nicotinate below those that can serve as sole nitrogen sources allow hypoxanthine utilization by \( hxA \) strains [16,17]. Nicotinate induces PHII, which catalyses the hydroxylation of hypoxanthine to xanthine. Xanthine is further hydroxylated to uric acid by a xanthine dioxygenase encoded by the \( xanA \) gene [18–20]. This is schematized in figure 1. The induction pattern implies that PHII belongs physiologically to the nicotinate utilization pathway and not to the purine utilization pathway.

In the 1970s and 1980s, we attempted to characterize genetically the nicotinate utilization pathway in \( A. nidulans \). The results have only been published schematically [1–3,22] and thus will be summarized below. We isolated mutants able to grow on hypoxanthine as a nitrogen source, but not on a medium that contains hypoxanthine, allopurinol and nicotinate (1 mM), which, at this concentration, does not serve as a nitrogen source but fully induces PHII [1]. The wild-type grows on this medium, as PHII (resistant to allopurinol inhibition [1,7]) hydroxylates hypoxanthine to xanthine, which is further hydroxylated to uric acid by the XanA protein (figure 1). Three groups of mutations, mapping in three different genes, were obtained. One group, \( hxnS \), PHII, results in the inability to grow on the isolation medium and on nicotinic acid as the sole nitrogen source (10 mM) but maintains its ability to grow on 6-OH nicotinic acid or a further metabolite in the nicotinate utilization pathway and under the control of the HxnR/AplA transcription factor(s). Full references are given in the text.

![Figure 1. Metabolic cross-talk between the purine and nicotinate utilization pathways. PHI is a conventional XDH able to catalyse the conversion of hypoxanthine to xanthine and xanthine to uric acid. XanA is an \( \alpha \)-ketoglutarate-dependent xanthine dioxygenase, accepting xanthine but not hypoxanthine as a substrate. From there uric acid is converted into ammonium (\( \text{NH}_4^+ \)) by the well-established purine utilization pathway ([21] for review). PHII is an unconventional MOCO carrying enzyme hydroxylating hypoxanthine to xanthine and nicotinic acid to presumably 6-OH nicotinic acid. As this latter compound is a nitrogen source, it is presumably converted into ammonium, which is indicated by a dashed blue connector. Note that unlike PHI, PHII cannot use xanthine as a substrate. In black: steps induced by uric acid, under the control of the UaY transcription factor. In blue: steps actually (\( hxnS \), PHII) or presumably induced by nicotinic acid, 6-OH nicotinic acid or a further metabolite in the nicotinate utilization pathway and under the control of the HxnR/AplA transcription factor(s). Full references are given in the text.](image)

**Table 1.** A summary of the properties of PHI and PHII compiled from the literature. Data from Lewis et al. [7] for the properties of the enzymes in crude extracts and from Mehra and Coughland [8] (PHI) and [11] (PHII) for the purified enzymes. The reader is referred to the original articles for further details. R. rate, relative rate to hypoxanthine, given an arbitrary value of 1. The concentration of each substrate was 2.5 times its \( Km \).

| substrate                        | PHI (HxA) | PHII (HxnS) |
|----------------------------------|-----------|-------------|
|                                  | crude extract | pur. enzyme | crude extract | pur. enzyme |
| hypoxanthine (6-hydroxypurine)   | 1.00      | 1.00        | 1.00          | 1.00        |
| xanthine (2,6-dihydroxypurine)   | 0.63      | 0.61        | 0.61          | 0.61        |
| 2-hydroxypurine                  | 0.59      | 0.49        | 0.49          | 0.49        |
| allopurinol (4-hydroxypyrazolo-3,4-dipirimidine) | <0.005 | 0.007      | 0.007         | 0.007       |
| nicotinate                       | —         | —           | 0.16          | 0.22        |

\( Km \)s of competitive inhibitors with hypoxanthine as a substrate.
mutations define the structural gene for PHII. Non-leaky \( hxnS \) mutations resulted in the loss of PHII enzyme activity but were heterogeneous regarding PHII cross-reacting material (CRM) [2,22]. Furthermore, mutations in \( hxnR \) result also in the complete inability to grow on 6-OH nicotinate. \( hxnR \) mutants are non-inducible for PHII activity or CRM [22]. The \( hxnR \) mutations are fully recessive and thus represent loss-of-function mutations. They define an activating transcription factor, necessary for the expression of \( hxnS \) and at least one other enzyme of the nicotinate utilization pathway, involved in the downstream conversion of 6-OH nicotinate. A number of mutants constitutive for PHII were called \( aplA' \) [1]. These represent regulatory gain-of-function mutations [1]. The \( aplA \) and \( hxnR \) mutations could represent two tightly linked genes or a single gene where the relatively frequent constitutive mutations define (a) negative-acting domain(s). The \( hxnS, hxnR \) and \( aplA' \) mutations are tightly linked on chromosome VI (less than 1 centiMorgan for crosses involving several alleles of the three classes). One mutation isolated, described elsewhere, defines the \( xanA \) gene [18,19]. We report here that \( hxnS \) and \( hxnR \) are part of an extended gene cluster that includes four additional co-regulated genes. The \( aplA' \) mutations map in specific domains of the \( hxnR \) gene product. We discuss the evolutionary relationships between the structurally similar but functionally distinct HxA and HxnS paralogues, the domain structure of HxnR and the conservation of the nicotinate gene cluster in the Aspergillaceae.

2. Results

2.1. Identification and characterization of the \( hxnS \) gene

We expected the \( hxnS \) gene to be a paralogue of \( hxA \) [7,16]. In the \( A. nidulans \) genomic sequences of the Cereon Aspergillus Sequencing Project (later incorporated into the Aspergillus Genome Database, AspGD [23]), we found an incomplete homologue of \( hxA \) [9]. We localized the sequence encoding this XDH paralogue to chromosome VI cosmid W31:H08 (see ‘Material and methods’ section), in line with the mapping of \( hxnS \). This cosmid complements both the \( hxnS41 \) and the \( hxnR2 \) loss-of-function mutations. We sequenced the region comprising the putative \( hxnS \) gene to reveal a protein with very high (51%) identity to PHI encoded by the \( hxA \) gene and identical with the protein specified by the AN9178 locus in the AspGD genome database (GenBank accession number KY962010). The cognate full-length cDNA sequence was also obtained (GenBank accession number KX858438). The \( hxnS \) open reading frame is interrupted by three introns in different positions to those extant in \( hxA \) (figure 2). The \( hxnS \) gene encodes a protein of 1396 residues (Hxa, 1363 residues). The molecular masses are compatible with those experimentally determined for PHI and PHII native dimers [7] and with the slower migration of HxnS seen in the electronic supplementary material, figure S1, in native polyacrylamide gels. We deleted the putative \( hxnS \) gene (see ‘Material and methods’ section). The deletion strain is able to grow on hypoxanthine, unable to use nicotinate as a nitrogen source and unable to grow on media containing hypoxanthine (N-source), allopurinol (inhibitor of PHI) and 100 \( \mu \)M nicotinate or 6-OH nicotinate (as inducer), which requires HxnS activity (figure 3; electronic supplementary material, figure S1, the latter showing enzyme activities with both hypoxanthine and nicotinic acid as substrates in native gels). \( hxnS\Delta \) strains are able to use 6-OH nicotinic acid as a nitrogen source, albeit at a reduced level (figure 3; electronic supplementary material, figure S1). The significance of the latter is not clear as the cognate parent strain also uses 6-OH nicotinate badly, and an \( hxB20 \) strain (see sections ‘Introduction’ and ‘A tightly co-regulated gene cluster in chromosome VI’, for HxB function) does not seem to be impaired in its utilization (figure 3). Previously isolated \( hxnS \) mutations result in the same phenotype as \( hxnS\Delta \) on the N-source hypoxanthine supplemented with allopurinol or on nicotinate. However, they do not show any impairment in 6-OH nicotinate utilization (electronic supplementary material, figure S1). The three classical loss-of-function mutations available were all isolated in an \( hxnR'/7 \) background, which results in overexpression of other genes under HxnR control (see below) encoding other proteins putatively involved in 6-OH nicotinate utilization (figure ba,b). The \( hxnS35 \) and \( hxnS41 \) alleles are consistent with those experimentally determined, while \( hxnS29 \) results in a Phe1213Ser change in a conserved region (figure 2). The \( hxnS35 \) and \( hxnS41 \) mutations result in loss of PHIII CRM, as assessed by immunoprecipitation, while \( hxnS29 \), a leaky mutation on allopurinol supplemented hypoxanthine medium (see electronic supplementary material, figure S1), fully retains CRM [22]. The above constitutes formal evidence that the locus AN9178 specifies the \( hxnS \) gene. Strains carrying the \( hxnS29 \) mutation have a clear phenotype in vivo, despite showing HxnS activity in vitro (electronic supplementary material, figure S1). The Phe1213Ser mutation may affect the stability rather than the activity of the enzyme. We have checked if 6-OH nicotinate (i.e. the product of nicotinate hydroxylase activity) could also be a substrate for HxnS. A very faint staining can be seen after 48 h incubation, a signal not stronger than the one obtained in the absence of substrate, incubating the gel in the presence of the tetrazolium salt (not shown).

2.2. A comparison of HxnS (PHII) with HxA (PHI)

Figure 2 compares PHI (HxA) and PHII (HxnS) to the thoroughly chemically and structurally characterized \( B. taurus \) XDH enzyme [24,27]. HxnS and HxA and their fungal orthologues (see below) differ less from each other than other eukaryotic XDH paralogues, such as so-called ‘aldehyde oxidases’ from genuine XDHs. Eukaryotic ‘aldehyde oxidases’, so denominated for historical reasons, are enzymes very similar to XDH, but with different substrate specificities [30,31]. Features that differentiate HxnS from HxA and those that are conserved in HxA and HxnS putative fungal orthologues are discussed below. The residues involved in the two amino-terminal 2Fe/2S clusters, and the FAD- and NAD-binding residues identified in the crystal structure of the \( B. taurus \) enzyme are strictly conserved in HxA and HxnS (figure 2). HxnS comprises several insertions when compared with HxA and other characterized XDHs (figure 2). The first insertion occurs between the second and the third Cys residues of the second 2Fe/2S cluster. The sequence between the 2Fe/2S cluster domain and the FAD/NAD-binding domain is longer in HxnS. Within the FAD/NAD domain, the residue corresponding to Phe417 of
the \( \text{B. tauru} \) \( \text{XDH} \) is almost universally an aromatic residue in \( \text{XDHs} \) (Tyr454 in \( \text{HxA} \)) but it is Ile (Ile478) in \( \text{HxnS} \) and always an aliphatic hydrophobic residue in \( \text{HxnS orthologues} \) (figure 2). The carboxy-terminal MOCO/substrate-binding domain (starting from residue 590 in the \( \text{B. tauru} \) \( \text{XDH} \)) shows an almost complete conservation of both the residues interacting with MOCO [25] and those interacting with substrates, including most of the residues that line the substrate access channel. His954 of the \( \text{B. tauru} \) enzyme, a residue not involved in the enzyme active site, is conserved in \( \text{HxA} \) (His985) and in most of its orthologues. However, it is Pro (Pro1008) in \( \text{HxnS} \) (figure 2) and in all its putative orthologues. This change does not affect the modelled secondary structure (not shown, but see below). Other amino acid residues, which differ systematically among \( \text{HxA} \) and \( \text{HxnS orthologues} \) (see section below), correspond to some of the residues involved in MOCO binding; the Val1081 and Ser1082 of the \( \text{B. tauru} \) enzyme are Ala1112 and Ser1113 in \( \text{HxA} \) but Ser1137 and Gly1138 in \( \text{HxnS} \) (figure 2). Conserved residues include Arg880 of \( \text{B. tauru} \) XDH (Arg911 of \( \text{HxA} \) and Arg934 of \( \text{HxnS} \)), a residue that is never conserved in XDH-like aldehyde oxidases [9,29–31]. Mutations affecting this residue in \( \text{hxA} \) result in altered substrate specificity including a PHII-like resistance to allopurinol inhibition and the inability to accept xanthine as a substrate [18,29]. Glu803 of the \( \text{B. tauru} \) enzyme is conserved in \( \text{HxA} \) (Glu833) and \( \text{HxnS} \) (Glu856). This key residue is never conserved in XDH-like aldehyde oxidases [30]. Within the \( \text{HxA} \) MOCO/substrate-binding domain, several mutations result in either loss-of-function or altered substrate specificity phenotypes [29]. All the corresponding residues involved are conserved in \( \text{HxnS} \) (figure 2). The pair of aromatic amino acids that sandwich the purine ring and orient the substrate towards the MOCO are conserved (Phe914 and 1009 in the \( \text{B. tauru} \) enzyme, 954 and 1040 in \( \text{HxA} \), 968 and 1064 in \( \text{HxnS} \)).

A striking exception to the sequence conservation is the insertion of an Ala ( Ala1065 in \( \text{HxnS} \)) between the almost universally conserved Phe1009 and Thr1010 (numeration as in [25], figure 2). The Phe/Thr pair is also conserved in bacterial XDHs (residues 459 and 460 in subunit B of the \( \text{Rhodobacter capsulatus} \) XDH [32]). An Ala insertion at this position is an almost absolute feature of \( \text{HxnS orthologues} \) (FATAL in \( \text{HxnS orthologues} \), FSTAL in \( \text{Choiromyces venosus} \) putative XDH).
The complete genotypes are given in the electronic supplementary material, figure S3 and table S1. Enzymes fungal genomes for homologues of XDH (see electronic supplementary material, HZS.135), hxnR compared with Thr1041 (HxA) and Thr1010 (HxnS) with the obvious exception of the orientation of Thr1066 of two separated clades of the XDHs of the Basidiomycota (one of them clustering with HxnS orthologues), the divergence of Saitoella complicata and the Taphrina spp., and the position of both the Saccharomycotina and Taphrinomycotina as outgroups of HxA orthologues rather than as an outgroup of all the Eurotiomycetes: orthologues of HxA are present in all species available, but the presence of HxnS is patchy, i.e. present in the nidulantes group and the black aspergilli, but not, for example, in A. flavus. With the exception of Penicillium paxilli and P. citrinum, which contain hxnS orthologues (unlinked to hxnR; see below), the hxnS orthologues are missing from species of Penicillium. Within the Sordariomycetes, a similar pattern of loss occurs, with hxnS orthologues present in the Nectriaceae (order Hypocreales), but not in the Sordariales (such as Neurospora crassa, Sordaria macrospora and Podospora anserina). The only PH-like enzyme present in O. maius (Leotiomycetes) could represent a second neo-functionalization, in which an enzyme phylogenetically related to HxnS would have reacquired HxA substrate specificity (see comments to electronic supplementary material, figure S3). The phylogeny (figure 4; electronic supplementary material, figure S3) strongly suggests a duplication of an HxA ancestral gene occurring at the root of the Dikarya. This duplication would have been followed by either neo-functionalization, leading to HxnS (in the Pezizomycotina) or loss of one of the two ancestral paralogues with HxA function (elsewhere in the Dikarya). This discrepancy between the timing of duplication and neo-functionalization would account for the two separated clades of the XDHs of the Basidiomycota (one of them clustering with HxnS orthologues), the divergence of Saitoella complicata and the Taphrina spp., and the position of both the Saccharomycotina and Taphrinomycotina as outgroups of HxA orthologues rather than as an outgroup of all the Pezizomycotina PFs (see electronic supplementary material, figure S3 legend).

The hxnS orthologues, which have been included in figure 4 and the electronic supplementary material, figure S3, show a highly variable exon–intron structure, as discussed in the supplementary material (comments on the exon–intron structure of hxnS orthologues).

2.4. Identification and characterization of the hxnR/aplA gene

Closely linked to, but separated by locus AN9177 (to be called hxnT; see below), there is an open reading frame of 2673 nt (interrupted by a single 75 nt intron) encoding a protein of 865 residues comprising two typical Cys2His2
Figure 4. A simplified phylogeny of the fungal purine hydroxylases, HxA (PHI)-like and HxnS (PHII)-like. This tree in cartoon form was extracted from the more complete tree shown in the electronic supplementary material, figure S3, where all species used are indicated. Outgroups are the nearest non-fungal taxa of the Opisthokonta. Values at nodes are aLRTs (approximate likelihood ratio tests). The arrows indicate the putative nodes where the gene duplication and the PHII neo-functionalization occurred.
Zn fingers near its amino terminus (AN11197). We have resequenced this region (GenBank accession number KX669266). In the *A. nidulans* open reading frame, there are two possible in-phase initiation codons separated by three residues (MKAKM; electronic supplementary material, figure S4). In other aspergilli available in the databases, only the second Met codon is present. As the first codon is not conserved, there could be more than one reading frame (RNAseq data, J-Browse module at http://www.aspgd.org/), we have assumed that this is the genuine start codon in *A. nidulans* (in accordance with Kozak [37]). Between residues 394 and 668, a PFAM domain ‘Fungal transcription specific domain’ PF04082 was detected (figure 5a). A nuclear localization signal from residues 77 to 87 (NLS, VLETKRMRRRA) downstream from the Zn fingers is strongly predicted by cNLS MAPPER, while a nuclear export signal (NES, LDIDL) is predicted for residues 285–289 by NerNES (figure 5a).

We deleted the whole AN11197 coding region. The resulting phenotype is identical to that reported previously for *hxnr* loss-of-function mutations [1,3,22] (figure 3; electronic supplementary material, figure S1 and transcriptional phenotypes in section ‘A tightly co-regulated gene cluster in chromosome VI’): inability to use nicotinate and 6-OH nicotinic acid as sole nitrogen sources, to which we can add now the inability to use 2,5-dihydroxypyridine, an intermediate in the catabolism of nicotinate in bacterial species [39,40]. Figure 3 confirms that 2,5-dihydroxypyridine is an inducing intermediate in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of *hxns*.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.

Extant loss of function, as well as constitutive mutations (*alpr* mutations; see ‘Introduction’ section) map within the *hxnr* open reading frame (figure 5a). We have thus renamed the constitutive regulatory mutations, *hxnr*†. We attempted to define the domain(s) comprising residues mutable to constitutive mutations in *A. nidulans* as this metabolite allows strong growth on hypoxanthine in the presence of allopurinol, which necessitates induction of 2,5-dihydroxypyridine.
a larger domain between residues 565 and 639. For a number of residues we have obtained several different amino acid changes. Accessible aromatic residues at positions 226 and 228 and a basic residue at position 605 seem necessary for HxnR to be in its default, inactive state, in the absence of its physiological inducer.

We detected putative HxnR orthologues only among the Pezizomycotina (electronic supplementary material, table S3). We would expect a strong correlation between the presence of hxnR and hxnS orthologues. Out of 139 species of the Pezizomycotina screened, 40 have only hxnR and 14 only hxnS (electronic supplementary material, table S4). Among the 85 species where both genes are extant, tight clustering is evident in most of them (see the section ‘Conservation of the hxn gene cluster in the Aspergillaceae’). The absence of clustering is common among the Sordariomycetes, with the exception of the Xylariales order where the clustering is maintained. These 85 species include all classes of the Pezizomycotina subphylum with the exception of the Orbiliomycetes and the Lecanoromycetes.

2.5. A tightly co-regulated gene cluster in chromosome VI

In A. nidulans, the hxnR and hxnS genes are within a cluster of co-regulated genes. This is shown in figures 6 and 7. Six neighbouring genes, inducible by nicotinate and 6-OH nicotinate, are non-inducible in strains carrying either the hxnR2 or hxnRΔ mutations and show strong constitutive expression in the hxnR7 background. The genes in the cluster are: hxnS (AN9178), hxnT (AN9177), hxnR (AN11197), hxnP (AN11189), hxnY (AN11188) and hxnZ (AN11196) (figure 6d). The flanking genes AN9179 (adjacent to hxnS) and AN9174 (adjacent to hxnZ and transcribed convergently) are not induced by nicotinate and they are not affected by hxnR constitutive or loss-of-function mutations (not shown). The hxnP and hxnZ genes encode transmembrane proteins of the Major Facilitator superfamily (PF07690.13). hxnT encodes a flavin oxidoreductase (Oxidored_FMN, PF00724), while hxnY encodes a typical α-ketoglutarate-dependent dioxygenase (PF14226.5 and PFU0171.19). The role of each gene in nicotinate utilization and their phylogenetic relationships will be discussed elsewhere, HxnP and HxnZ being involved in the uptake of nicotinate-derived metabolites, and HxnT and HxnY in the further metabolism of 6-OH nicotinic acid (E Bokor, M Flipphi, J Ámon, C Sczaczecchio and Z Hamari, unpublished results). We can however state that, for each of these genes, the nearest homologue is a fungal and not a bacterial gene (not shown). hxnR is itself an inducible gene (figure 6a,b).

There is a clearly detectable level of hxnR transcript under non-induced conditions, at variance with the other genes of the cluster. RNASeq data [23,42], available in J Browse (http://www.aspgd.org/), confirm the co-regulation of the cluster, where all genes in this cluster are non-expressed in conditions of nitrogen sufficiency and derepressed by nitrogen starvation. Under our experimental conditions, with the exception of hxnR, genes in the cluster are virtually non-expressed in media that contain good nitrogen sources but are expressed under nitrogen-starved conditions (figure 6c). All genes in the cluster are drastically repressed by ammonium (figures 6b and 7a). HxnR is necessary for expression under nitrogen-starved conditions (figure 6c).

The strong constitutivity of hxnRΔ7 strains is clear in the presence of a non-repressive nitrogen source (acetamide) or under conditions of nitrogen starvation. The transcript of hxnR, which had previously been found to be independently regulated by HxnR and UaY (and thus independently induced by nicotinate and uric acid [41]) behaves qualitatively as the five structural genes in the cluster (figure 6b).

The role of AreA, the GATA factor mediating nitrogen metabolite derepression [43–45], is shown for hxnS and hxnP in figure 7a. Transcription of both hxnS and hxnP is abolished in a strain carrying a null areA mutation (areA6000) under all conditions, including nitrogen starvation. Surprisingly, the transcription of both hxnS and hxnP is diminished in a strain carrying an xprD1 mutation (considered to be the most extremely derepressed allele of areA, ([46] and references therein)); the allele is called xprD1 for historical reasons [43,47]. By contrast with the genes of the nitrate and purine assimilation pathways [48–50], the glutamate–aspartate transporter gene agtA [51] and also xrbB [41], the hxnS and hxnP genes are fully repressed by 10 mM ammonium in an xprD1 strain. A similar atypical effect has been reported for the main urea transporter ureA gene [52].

A downstream metabolite of nicotinate is the physiological inducer of the HxnS protein [12]. Figure 7b shows this to be the case at the level of mRNA steady-state levels for both hxnS and hxnP. In an xrbB null mutant lacking HxnX activity, nicotinate does not behave as an inducer but 6-OH nicotinate does. Thus the effector of HxnR is not nicotinate but 6-OH nicotinate or a metabolite further downstream the nicotinate utilization pathway. The in vivo test shown in figure 3, where 2,5-dihydroxypteridine acts as inducer, suggests the latter to be the case.

2.6. Conservation of the hxn gene cluster in the Aspergillaceae

The evolution of the whole nicotinate utilization pathway in fungi will be dealt with in another publication (E Bokor, M Flipphi, J Ámon, C Sczaczecchio and Z Hamari, unpublished results), but we discuss here the conservation of the hxn cluster in the Aspergillaceae family. Examples of the organization of the cluster are shown in figure 8. Episodes of gene gain and loss are shown, including the duplication of hxnY or hxnT as well as the loss of hxnT and hxnS. In A. ochraceoroseus, only hxnS is present, a mirror image of the situation in A. flavus (and other species in section Flavi) and P. digitatum (and all other Penicillium species but two), where the genome includes all hxn genes with the exception of hxnS. The absence of hxnS may imply that, in these species, the cluster deals with the utilization of nicotinate derivatives (such as 6-OH nicotinic acid) rather than that of nicotinate per se.

The situation in P. citrinum (and P. paxilli) implies a secondary reconstitution of the pathway by horizontal transmission, as an unlinked orthologue of hxnS is present, seemingly reacquired from a member of the Hypocreales (order of Sordariomycetes; see Phylogeny section, electronic supplementary material, figure S3). In A. niger, and related black aspergilli albeit not in A. carbonarius or A. aculeatus, a new, intronless gene, encoding a putative nitroreductase is inserted in the cluster, between hxnR and hxnT (see ‘Discussion’ section).
3. Discussion

3.1. A nicotinate-inducible eukaryotic cluster

With the exception of some of our own old work (see ‘Introduction’ section) no genes or enzymes involved in the degradation of nicotinate have been described in any eukaryote. Degradation of nicotinic acid has been studied in plant cell cultures and tea plant material fed with carboxyl-C14-nicotinic acid and C14-6-nicotinic acid, monitoring the formation of 14CO2 [53–55]. No enzymes involved in these processes were identified and inspection of relevant genomes only revealed one typical XDH. Thus, the A. nidulans HxnS is the hitherto only eukaryotic nicotinate hydroxylase studied, and the hxn gene cluster we have identified is the first co-regulated eukaryotic gene cluster involved in the utilization of nicotinate ever described.

3.2. The hxA/hxnS duplication compared to other eukaryotic MOCO-enzyme duplications

Neo-functionalization of enzymes of the XDH group arising from ascertained or presumed gene duplications occur in...
both prokaryotes and eukaryotes. Fetzner and co-workers have described the diversity of bacterial MOCO enzymes of the XDH group, even if the phylogeny of these enzymes with different specificities remains unstudied [39,56–59]. Duplication and neo-functionalization of genes encoding XDH-like enzymes are widespread in Metazoa, studied mainly in insects and vertebrates. In Metazoa, a close linkage of the neo-functionalized genes with strict conservation of intron/exon structure is the rule [30,60–63]. The implication is that XDH gene duplication has occurred by different mechanisms in metazoans and fungi. In metazoans, duplication seems to occur at the DNA level by unequal crossover. In the fungi, the striking amino acid sequence conservation among the HxA/HxnS paralogues together with the variability of intron positions suggests that the duplication of an HxA ancestral gene occurred via retroposition, followed by a re-intronization either after or concurrent with re-functionalization of the duplicated gene. Notwithstanding the mechanism underlying this gene duplication, the hxA/hxnS duplication is quite ancient, occurring before the divergence of the Taphrinomycotina from other Ascomycota (greater than 400 Ma [64,65]), which allows the possibility of intron loss and reinsertion. The variation of intron—exon organization in both the hxA and hxnS clades (not shown) is also consistent with this possibility (figure 4 and electronic supplementary material, figure S3, for the relevant positions in the phylogenetic tree).

3.3. Convergent evolution of bacterial and fungal nicotinate hydroxylases

MOCO enzymes able to catalyse the hydroxylation of nicotinate to 6-OH nicotinate have been described in a variety of bacterial species [40,66–68]. However, it can be excluded that the HxnS proteins have originated by horizontal transmission from bacteria. In all eukaryotes, XDH-like enzymes are dimers of chains of approximately 1500 amino acid residues comprising three discrete domains (figure 2). In bacteria, these domains are encoded by at least two genes, one specifying a small subunit carrying the 2Fe/2S centres and the FAD-binding sequences and a large subunit carrying the MOCO and substrate-binding centres. These genes, included in an operon, reflect the amino- to carboxy-terminus order of the domains in the eukaryotic XDH-like enzymes. A similar structure occurs in bacterial nicotinate hydroxylases [40,67], which makes improbable a direct bacterial origin of the fungal nicotinate hydroxylases. Figure 4 and the electronic supplementary material, figure S3 show that HxnS orthologues have originated by gene duplication within the fungal kingdom, possibly at the root of the Dikarya, with a neo-functionalization process occurring within the Pezizomycotina subphylum. BLASTP screening with the MOCO/substrate-binding domains of both HxA and HxnS (figure 2) against all bacterial sequences available in the NCBI non-redundant protein (nr/nt) database yielded homologues of the MOCO-binding subunits of putative bacterial XDHs, but in no case (among the first 100 sequences) were MOCO subunits of known bacterial nicotinate hydroxylases found (not shown). Thus, fungal nicotinate dehydrogenases show more similarity to bacterial XDHs (and as a matter of course, to all genuine eukaryotic XDHs) than to bacterial nicotinate dehydrogenases.

A comparison of the sequences of the MOCO and substrate-binding subunits of bacteria suggests that there are at least three classes of MOCO nicotinate hydroxylases, exemplified by Pseudomonas putida, Eubacteriumarkeri and Bacillus niacinis, respectively [40,66,67]. We have stated that the insertion of an Ala residue (HxnS Ala1065) between the conserved Phe and Thr in the active site is a signature of fungal HxnS orthologues (see §§2 and 3 of Results). Genuine
bacterial XDHs also carry an FT motif in the cognate place in the structure [32,34]; however, the divergence of bacterial nicotinate hydroxylases from bacterial XDHs is such that neither sequence alignments nor structural modelling (not shown) gave a clear inking of which modifications resulted in, or at least correlate with the shift in substrate specificity. It seems that not only has there been convergent evolution of fungal and bacterial nicotinate hydroxylases, but that nicotinate hydroxylases evolved several times independently within bacteria. A hint to the shift in specificity towards hydroxylation of nicotinic acid is provided by the molecular structure of the nicotinate hydroxylase of *E. barkeri* [68,69].

In this enzyme, the substrate/MOCO-binding domain is split into two independent peptides (L and M). Strikingly, it carries a selenium rather than a sulfur atom as the terminal ligand to the Mo(VI). Selenium also occurs in the XDH of this organism [68,69], which is consistent with an independent evolution of the *E. barkeri* nicotinate hydroxylase from that of other bacterial enzymes of similar specificity (see above). Most active-site residues are conserved, with an interesting exception. Tyr13 of subunit M (Tyr13_M) is modelled to hydrogen-bind the heterocyclic N atom of nicotinate by its hydroxyl group [68]. The corresponding residue in the *B. taurus* XDH is Phe1005, which has not been proposed to interact with the substrate [24,27]. This Phe residue, four residues upstream of Phe1009 (of *B. taurus*), is conserved in both HxA and HxnS (figure 2) and indeed in all HxA and HxnS-like fungal enzymes included in the electronic supplementary material, figure S3, with the exception of the putative XDH of the four divergent *Taphrina* species, where it is substituted by a His. Tyr13_M is not conserved in several other characterized or putative nicotinic acid hydroxases such as those of *Ps. putida* and its putative orthologues, where the corresponding residue is an Arg or a His. No sequence similar to FTAL or FATAL is present in the enzyme of *E. barkeri* and its putative orthologues (see fig. S1 of [68]). It is tempting to speculate that the change in orientation of Thr1066 in HxnS allows an interaction with nicotinate by its hydroxyl group similar to that seen for modelled Tyr13_M in *E. barkeri*. Biochemical evidence indicates that the carboxyl group of nicotinate is essential for substrate binding of HxnS [11]; the hydroxyl of Thr1006 could potentially hydrogen-bind the carboxyl group of nicotinic acid. Differently from HxnS, the enzyme of *E. barkeri* does not accept hypoxanthine as substrate [68]. It can be proposed that the bacterial enzymes (at least the *E. barkeri* one) have fully evolved into dedicated nicotinate hydroxylases, while the HxnS orthologues conserve properties of XDH. The specific situation discussed for the homologue of *O. nius*, which can be proposed to have reverted to a typical XDH activity from an HxnS-like enzyme (see electronic supplementary material, figure S3 legend), would be in line with this speculation.

### 3.4. An unusual specific transcription factor

Fungal transcription factors regulating specific metabolic primary or secondary pathways are generally of the Zn2Cys6 (zinc cluster) class, while Cys2His2 (zinc finger) factors are usually, with very few recorded possible exceptions [70,71], broad domain regulators of either metabolism and/or morphology. The closest characterized transcription factor that shares architecture and has sequence similarity with HxnR is Klfp of *Schizosaccharomyces pombe*. This factor is necessary for maintenance of long-term quiescence and its absence results in abnormal cell morphology in the quiescent state [72]. The nearest homologue and possible orthologue of Klfp in *A. nidulans* is the protein of unknown function encoded by AN6733. The latter is strictly conserved in a syntenic position in all aspergilli included in the AspGD database and putative orthologues are present in all sequenced members of the Pezizomycotina (not shown). As hxnR is only present in the Pezizomyccotina, it is tempting...
to speculate that it originated from a duplication of the possibly essential ancestral orthologue of AN6733, the duplicated gene being then recruited into the nicotinate utilization pathway.

The apparent high frequency of constitutive, gain-of-function mutations, and their mapping indicate that the HxnR protein is, in the absence of inducer, in a default state non-competent to elicit transcription, and that the domains where constitutive mutations map are instrumental in maintaining HxnR in this ‘closed’, inactive state. The amino-terminal cluster of constitutive mutations maps outside the PF04082 domain, in sequences that are conserved only among HxnR orthologues. The carboxy-terminal mutations map within the PF04082 domain conserved in Klf1, AN6733 and NCU05242 (the *N. cressa* orthologue of AN6733). Note, for example, mutations affecting Lys603 in HxnR, a residue conserved in these four proteins (figure 5; electronic supplementary material, figure S4). The PF04082 domain of GalHp (244–537) coincides with the central regulatory domain of similarity proposed by Poch [73]; see also Stone & Sadowski [74]. The cognate domain of the *A. nidulans* NirA (pathway-specific regulation of nitrate assimilation) spans residues 230–487 [75]. Within this region maps a cryo-sensitive, non-inducible mutation (Arg347Ser) as well as its intragenic suppressors, some of which result in constitutivity. This domain possibly interacts (Arg347Ser) as well as its intragenic suppressors, some of which result in constitutivity. This domain possibly interacts.

### 3.5. The evolution of clustering

Old genetic and newly acquired data, which will be reported elsewhere (E Bokor, M Flipphi, J Amon, C Scazoncio and Z Hamari, unpublished data), established that not all the genes involved in nicotinate catabolism are within the *hxnZ–hxnS* gene cluster. We have described the conservation of this cluster within the Aspergillaceae. We discussed the origin of both *hxnS* and *hxnR* within the Pezizomycotina subphylum. While the selective pressures that led to the conservation of clustering of genes of a specific metabolic pathway have been the subject of animated discussion [76–78], we have no inkling of the recombination processes that led to clustering of the *hxn* genes in the first place. A model of recent local gene duplication can be excluded for the origin of all genes in the cluster, each nearest paralogue in the same organism being in every case unlinked and actually on a different chromosome (not shown). Within the Aspergillaceae, *A. nidulans* represents the possible primeval situation, with a pattern of both loss and duplication for other members of this family (figure 8). Recent duplication has occurred for some of the genes in the cluster. In *Monascus* sp. (exemplified by *M. ruber* in figure 8) an unlinked parologue of *hxnT* is extant, showing 58% amino acid identity with the copy within the cluster and a strict conservation of intron positions. Duplicated paralogues of *hxnY* occur in the *flavit* group and in species of the section Aspergillus. The fact that these duplicated genes are unlinked to the cluster excludes a model of duplication by unequal crossover.

It is noteworthy that instances of duplications are coupled with instances of loss. Duplication of *hxnY* in the section Aspergillus (exemplified by *A. ruber*, figure 8) is coupled with the loss of *hxnT*, while that of *hxnT* in section Flavi (exemplified by *A. flavus*, figure 8) is coupled with the loss of *hxnS*. This coupling may result from just one single recombination event. Note that in *M. ruber*, where there is no gene loss, duplication of *hxnT* coincides with the separation of *hxnS* from the cluster. The duplication of *hxnY*, with conservation of (some) intron positions, seems to have occurred before the divergence of the _flavi_ and the _fumigati_ groups. Remarkably, only the duplicated *hxnY* paralogue is retained in *A. fumigatus* and *Neosartorya fischeri*.

Horizontal transmission from pre-existent clusters has been established for both primary and secondary metabolism pathways. It has been proposed that nitrate assimilation gene cluster of fungi was horizontally transmitted from oomycetes [79]. We can exclude such horizontal transmission as the origin of the *hxn* cluster. The nearest paralogue of all the genes comprising the cluster is another fungal gene, usually in the same organism (data to be presented elsewhere, E Bokor, M Flipphi, J Amon, C Scazoncio and Z Hamari, unpublished results).

One exception to this is the incorporation of an intronless nitroreductase gene into the *hxn* cluster of most aspergilli of the section *nigri* and its presence outside the cluster in four other aspergilli including *A. nidulans*. A phylogenetic analysis (not shown) establishes that this gene originates from a horizontal transfer from a cyanobacterium to an ancestral member of the Leotiomycteta (42% and 41% identity shared with the enzymes from *A. niger* and *A. nidulans*, respectively, with *nfsA* product from *Anabaena variabilis*; see [80] for a comparison of fungal and bacterial nitroreductases). Its incorporation within the *hxn* cluster of some aspergilli is quite intriguing. It may be relevant that many nitroreductases are involved in the degradation of N-heterocyclic compounds [81].

We have only presented a detailed phylogenetic analysis for *HxA/Hxns*, but work to be detailed elsewhere (E Bokor, M Flipphi, J Amon, C Scazoncio and Z Hamari, unpublished results) suggests that, with the one exception mentioned, all genes in the *hxn* cluster have originated from duplications within the Pezizomycotina, and that clustering followed or was synchronous with duplication. Similar evolutionary patterns for the clusters were described in fungi. A pattern of gene duplication and clustering underlies the origin and variable arrangement of the *alc* (ethanol utilization) gene cluster in the aspergilli [82]. These patterns of gene clustering resemble those described in plants, where genes organized in clusters involved in secondary metabolism originate from duplication of non-clustered genes of primary metabolism ([83–85] and references therein).

### 4. Material and methods

#### 4.1. Strains, media and growth conditions

The *A. nidulans* strains used and/or constructed in this work are listed in the electronic supplementary material, table S5. Standard genetic markers are described in http://www.fgsc.net/Aspergillus/gene_list/. Complete (CM) and
minimal media (MM) contained glucose as the carbon source; MMs supplemented with different N-sources were used [13,86]. The media were supplemented according to the requirements of each auxotrophic strain (www.fgsc.net). Nitrogen sources, inducers, repressors and inhibitors were used at the following concentrations: 10 mM sodium nitrate, 10 mM nicotinate (1 : 10 dilution from 1 M nicotinic acid dissolved in 1 M sodium hydroxide), 10 mM 6-OH nicotinic acid (1 : 100 dilution from 1 M 6-OH nicotinic acid dissolved in 1 M sodium hydroxide), 10 mM 2,5-dihydroxypropyridine, 1 mM hypoxanthine, 5 mM L(-)-diammonium-tartrate, 5 mM urrea, 1 mM acetamide as sole N-sources; 1 mM or 100 μM nicotinic acid sodium salt, 1 mM or 100 μM 6-OH nicotinic acid, 100 μM 2,5-dihydroxypropyridine and 0.6 mM uric acid as inducers; 5 mM L(-)-diammonium-tartrate as repressor; 5.5 μM allopurinol as inhibitor of purine hydroxylase I (encoded by hxA) enzyme activity. Cesium chloride at a 12.5 mM final concentration was used in mutagenesis experiments to reduce the background growth of the nitrogen-source non-utilizer strains (http://www.fgsc.net/Aspergillus/gene_list/supplement.html#other). The strains were maintained on CM; otherwise MM with various N-sources were used in the experiments supplemented with the required vitamins. The mycelia for protein extraction were grown for 14 h at 37°C shaken at 150 r.p.m. in MM with acetamide or urea as nitrogen sources and induced when appropriate after 12 h of growth with 6-OH nicotinate. For mRNA extraction, mycelia was grown on acetamide, or urea N-sources were used for growth for 10 h at 37°C with 150 r.p.m. and after 8 h of growth, nicotinic acid, 6-OH nicotinic acid or uric acid was added to the medium as inducer and ammonium as repressor. For total DNA extraction, mycelia were grown in MM with nitrate as a N-source.

4.2. Mutagenesis

For UV mutagenesis, 10⁹ conidia of A. nidulans strains HZS.98, HZS.248 and HZS.418 in 20 ml 0.01% Tween (in a Petri dish with a 14.5 cm diameter) were exposed to UV light (Philips TUV15 W 9L1, 254 nm) with gentle shaking (50 r.p.m.) for 20 min, resulting in 95% kill. For 4-nitroquinoline-1-oxide (4-NQO) mutagenesis, conidia of HZS.248 were mutagenized as previously described [87]. Spores were plated on MM with hypoxanthine as the sole nitrogen source supplemented with 5.5 μM allopurinol and 12.5 mM cesium chloride. Spores able to grow on this medium were expected to be hxnR constitutive (hxnR⁺) mutants. The presence of allopurinol resulted in the complete inhibition of purine hydroxylase I (encoded by hxA) in a recipient hxA⁺ strain (HZS.98), therefore the hypoxanthine utilization must result from the activity of purine hydroxylase II (encoded by hxnS), which requires either induction by nicotinate or 6-OH nicotinate or the presence of a constitutive mutation in the hxnR gene. In the hxA⁺ strain HZS.98, gain-of-function allopurinol-resistant mutations at the hxA locus also may occur. The hxnR⁺ hxA-linked allopurinol-resistant mutants, however, show reduced growth on hypoxanthine compared to hxA⁺ hxnR⁺ mutants [1,29].

4.3. Staining for enzyme activity in gels

Crude protein samples of mycelia were obtained from 300 ml liquid cultures incubated at 37°C with 180 r.p.m. agitation for 20 h, and induced after 15 h of growth with inducers where appropriate. Protein extraction was carried out as previously described [88]. The concentrations of crude protein samples were determined by the Bradford assay [89]. Native 10% PAGE using 0.025 M Tris, 0.19 M glycine cathode buffer (pH 8.3) according to Laemmli [90] was used to fractionate the crude extracts, containing 50 μg of protein/well. HxA- and HxnS-specific activities were detected by staining with hypoxanthine-tetrazolium [1], nicotinate-tetrazolium (100 mM pyrophosphate (pH 9.4), 1.27 mg ml⁻¹ iodonitrotetrazolium chloride and 0.5 mg ml⁻¹ nicotinic acid), while the diaphorase activity was detected with NADH-tetrazolium [16,91].

4.4. DNA and RNA manipulations

Total DNA was prepared from A. nidulans as described by Specht et al. [92]. For Southern blots [93] hybrid-N membranes (Amersham/GE Healthcare) were used and hybridizations were done by DIG DNA Labeling and Detection Kit (Roche) according to the manufacturer’s instructions. Transformations of A. nidulans protoplasts were done as described by Karacsony et al. [88] using a 4% solution of Glucanex (Novozymes, Switzerland) in 0.7 M KCl. For cloning procedures, Escherichia coli JM109 [94] and KS272 [95] were used and transformation of Es. coli was performed according to Hanahan [96]. Plasmid extraction from Es. coli and other DNA manipulations were done as described by Sambrook et al. [93]. Total RNA was isolated using a Nucleospin RNA Plant Kit (Macherey-Nagel) and RNase-Free DNase (Qiagen) according to the manufacturer’s instructions. cDNA synthesis was carried out with a mixture of oligo-dT and random primers using a RevertAid First Strand cDNA Synthesis Kit (Fermentas). Quantitative PCR (qPCR) and quantitative RT-PCR (qRT-PCR) were carried out in a CFX96 Real Time PCR System (BioRad) with SYBR Green/Fluorescein qPCR Master Mix (Fermentas) reaction mixture (94°C 3 min followed by 40 cycles of 94°C 15 s and 60°C 1 min). Specific primers were listed in the electronic supplementary material, table S6. Data processing was done by the standard curve method [97]. Northern blot analysis was performed using the glyoxal method [93]. In northern blots, equal RNA loading was calculated by optical density measurements (260/280 nm). [32P]-dCTP labelled gene-specific DNA molecules were used as gene probes using the random hexanucleotide-primer kit following the supplier’s instructions (Roche Applied Science). DNA sequencing was done by the Sanger sequencing service of LGC (http://www.lgcgroup.com). Primers used are listed in the electronic supplementary material, table S6.

4.5. Gene deletions

Deletion of hxnR and hxnS was obtained by Chaveroce’s method [95], which uses phage λ Red expressing Es. coli strain KS272 for obtaining the gene replacement by introducing a plasmid carrying the candidate gene and a PCR product of a transformation marker gene flanked with 50 bp regions of homology with the target DNA into the Es. coli strain (for details see the electronic supplementary material, Supplementary materials and methods). The obtained recombinant plasmid is then used for A. nidulans transformation in order to obtain an allelic exchange between the mutant allele on the plasmid and the wild-type locus. The detailed procedure is written in the electronic supplementary
material, Supplementary materials and methods. The first available \( hxnS\Delta \) strain (HZS.106 and its progeny HZS.254 used in the enzyme assays) was unfortunately found to carry additional ectopic copies of the recombinant plasmid, therefore a new \( hxnS\Delta \) strain (HZS.599) was obtained by the gene substitution method using the double-joint PCR [98] for constructing the gene substitution cassette (see electronic supplementary material, Supplementary materials and methods). All the genetic work and growth tests were done with the new \( hxnS\Delta \) strain.

4.6. In silico analysis

Sequence searches were carried out in both general (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and specialized databases (http://www.aspgd.org/, http://genome.jgi-psf.org/programs/fungi/index.jsf). We used (with permission) 59 unpublished DNA sequences from the JGI databases; the species involved are tagged with ‘*’ in the electronic supplementary material, table S1 footnote for further details). In every case, the gene models were manually derived by ourselves. Alignments were carried out with MAFFT (MAFFT E-INS-i and MAFFT G-INS-i); colour labelling of phylogeny was carried out with BMGE 1.0 (http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::BMGE) [99] and for phylogeny was carried out with BMGE 1.0 (http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::BMGE) [99] and maximum-likelihood phylogeny with PhyML 3.0 with automatic model selection (LG substitution model selected) [100,101] indicating approximate likelihood ratio tests [102]. Tree drawing was done with FIGTREE (http://tree.bio.ed.ac.uk/software/figtree/, http://mafft.cbrc.jp/alignment/server/) and localization signals were searched for at http://www.cbs.dtu.dk/services/TargetP/ [103], http://www.peroxisomedb.org/ [104], http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi [105], http://wolfpsort.org/ [106], http://genome.unmc.edu/ngLOC/cite.html [107]. Structural analysis and modelling was carried out with SWISS-PDBVIEWER [108] and I-TASSER, (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) [109,110], and model rendering with VMD 1.9. (http://www.ks.uiuc.edu/Research/vmd/) [111]. Structure superposition was done with the MultiSeq version integrated in VMD [112].

4.7. Statistical analysis

The significance of differences between datasets was determined by an unpaired \( t \)-test using the GraphPad Pssm 6 software.

Data accessibility. The datasets supporting this article are included in the paper and detailed in the electronic supplementary material tables. Sequences determined by us are available on GenBank (KY962010, KX585436, KX669266).

Authors’ contributions. Z.H. and C.S. conceived the project. J.M.K. and C.S. did the early genetic work. Z.H., J.A., R.F.-M., E.B. and A.C. carried out the molecular laboratory work. C.S. and M.F. did the phylogenetic analysis. Z.H., C.S. and M.F. wrote the manuscript. All the authors analysed the results and gave their final approval for publication.

Competing interests. We have no competing interests.

Funding. Work at Szeged was supported by the Hungarian National Office for Research and Technology (OTKA-K 101218), the Hungarian National Research, Development and Innovation Office (NKFKi-16 119516) and by the project GINOP-2.3.2-15-2016-00012. Work at Colchester was supported by the Science Research Council; work at Orsay was supported by the European Union HPRN-CT-1999-00084 (XONET), which also provided a studentship to A.C. and a fellowship to R.F.-M.; and Z.H. R.F.-M. and Z.H. at Orsay were supported by a Marie Curie Fellowship of the European Union (MCFI-2001-01804 and QLK4-CT-2002-51946).

Acknowledgements. C.S. thanks César Millan-Pacheco for insights on the use of VMD. JGI sequences used in the construction of the phylogenetic tree and the Consurf alignment were from the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/) in collaboration with the user community. We thank I. V. Grigoriev, F. Martin and J. Spatafora for permitting the use of genome sequences included in the 1000 Fungal Genomes project, the Metagenomics of soil fungi and the Mycorrhizal Initiatives. We thank J. Spatafora, F. Lutzoni, K. H. Wolfe, L. Connell, D. Armaleo, P. Dyer, S. Goodwin, A. Tsang, D. L. Nuss and A. Grum Grzhimailo for allowing access to the genomes of some individual species prior to publication.

References

1. Scazzocchio C, Holl FB, Foguelman AI. 1973 The genetic control of molybdoflavoproteins in Aspergillus nidulans. Allopurinol-resistant mutants constitutive for xanthine-dehydrogenase. Eur. J. Biochem. 36, 428–445. (doi:10.1111/j.1432-1033.1973.tb02928.x)

2. Scazzocchio C. 1980 The genetics of the molybdenum containing enzymes. In Molybdenum and molybdenum-containing enzymes (ed. MP Coughlan), pp. 487–515. Oxford, UK: Pergamon Press.

3. Scazzocchio C. 1994 The pure degradation pathway, genetics, biochemistry and regulation. Prog. Ind. Microbiol. 29, 221–257.

4. Pateman JA, Cove DJ, Raver BM, Roberts DB. 1964 A common co-factor for nitrate reductase and xanthine dehydrogenase which also regulates the synthesis of nitrate reductase. Nature 201, 58–60. (doi:10.1038/201058a0)

5. Schwarz G, Mendel RR, Ribbe MW. 2009 Molybdenum cofactors, enzymes and pathways. Nature 460, 839–847. (doi:10.1038/nature08302)

6. Amrani L, Primus J, Glatigny A, Arcangeli L, Scazzocchio C, Finnerty V. 2000 Comparison of the sequences of the Aspergillus nidulans hxA and Drosophila melanogaster ma-l genes with nifS from Azotobacter vinelandii suggests a mechanism for the insertion of the terminal sulphur atom in the molybdenopterin cofactor. Mol. Microbiol. 38, 114–125. (doi:10.1046/j.1365-2958.2000.02119.x)

7. Lewis NJ, Hurt P, Sealy-Lewis HM, Scazzocchio C. 1978 The genetic control of the molybdoflavoproteins in Aspergillus nidulans. IV. A comparison between purine hydroxylase I and II. Eur. J. Biochem. 91, 311–316. (doi:10.1111/j.1432-1033.1978.tb0967.x)

8. Mehna RK, Coughlan MP. 1989 Characterization of purine hydroxylase I from Aspergillus nidulans. J. Gen. Microbiol. 135, 273–278. (doi:10.1099/00221287-135-2-273)

9. Glatigny A, Scazzocchio C. 1995 Cloning and molecular characterization of hxA, the gene coding for the xantine dehydrogenase (purine hydroxylase I) of Aspergillus nidulans. J. Biol. Chem. 270, 3534–3550. (doi:10.1074/jbc.270.8.3534)

10. Coughlan MP, Mehna RK, Barber MJ, Siegel LM. 1984 Optical and electron paramagnetic resonance spectroscopic studies on purine hydroxylase II from Aspergillus nidulans. Arch. Biochem. Biophys. 229, 596–603. (doi:10.1016/0003-9861(84)90192-9)

11. Mehna RK, Coughlan MP. 1984 Purification and properties of purine hydroxylase II from Aspergillus nidulans. Arch. Biochem. Biophys. 229, 585–595. (doi:10.1016/0003-9861(84)90191-7)
niaD genes of Aspergillus nidulans contains multiple NitA binding sites which act bidirectionally. Mol. Cell Biol. 15, 5688 – 5699. (doi:10.1128/MCB.15.10.5688)

50. Cultrone A, Dominguez YR, Drevet C, Scaccozio C, Fernandez-Martín R. 2007 The tightly regulated promoter of the xanU gene of Aspergillus nidulans is included in a helixon. Mol. Microbiol. 63, 1577 – 1587. (doi:10.1111/j.1365-2958.2005.05609.x)

51. Apostolaki A, Erpapazoglou Z, Harispe L, Billimi M, Kafasa P, Kizis D, Penalva MA, Scaccozio C, Sophianopoulos V. 2009 AgTA, the dicarboxylic acid transport system of Aspergillus nidulans, is concertedly down-regulated by exquisite sensitivity to nitrogen metabolite repression and ammonium-elicited endocytosis. Eukaryot. Cell 8, 339 – 352. (doi:10.1128/EC.00270-08)

52. Abreu C, Sanguinetti M, Amillis S, Ramon A. 2010 Urea, the major urea/H+ symporter in Aspergillus nidulans. Fungal Genet. Biol. 47, 1023 – 1033. (doi:10.1016/j.fgb.2007.07.004)

53. Willeke U, Heeger V, Neuhann H, Schindelmüller I, Vordenfeld K, Barz W. 1979 Mutually exclusive occurrence and metabolism of trigonelline and nicotinic acid arabinoside in plant cell cultures. Phytochemistry 18, 105 – 110. (doi:10.1016/S0031-9422(00)90924-5)

54. Komossa D, Barz W. 1988 Glutaric acid as a substrate specificities. Cell Mol. Life Sci. 70, 1807 – 1830. (doi:10.1007/s00018-012-1229-5)

55. Terao M, Komoski M, Demontis S, Zanotta S, Garattini E. 1998 Isolation and characterization of the human aldheyde oxidase gene: evidence of intron/exon boundaries with the xanthine oxidoreductase gene indicates a common origin. Biochem. J. 332, 383 – 393. (doi:10.1042/ bj3320383)

56. Lucking R, Huhndorf S, Pfister DH, Plata ER, Lumbusch HT. 2009 Fungi evolved right on track. Mycologia 101, 810 – 822. (doi:10.3852/09-016)

57. Stajich J, Bebee ML, Blackwell M, Hickett DS, James TY, Spatafora JW, Taylor JW. 2009 The fungi. Curr. Biol. 19, R804 – R845. (doi:10.1016/j.cub.2009.07.004)

58. Nagel M, Andreassen JR. 1990 Purification and characterization of the molybdenzymes nicotinamide dehydrogenase and 6-hydroxynicotinamide dehydrogenase from Bacillus ninaici. Arch. Microbiol. 154, 605 – 613. (doi:10.1007/BF02248844)

59. Alhapel A, Darley DJ, Wagener N, Ekel M, Elnser N, Pirkel AJ. 2006 Molecular and functional analysis of nicotinamide catabolism in Eubacterium barkeri. Proc. Natl Acad. Sci. USA 103, 12341 – 12346. (doi:10.1073/pnas.0601635103)

60. Wagener N, Pirkel AJ, Ibdah A, Hille R, Dobberh V. 2009 The Mo-So active site of nicotinamide dehydrogenase. Proc. Natl Acad. Sci. USA 106, 11 055 – 11 060. (doi:10.1073/pnas.0902210106)

61. Schrader T, Rienhofer A, Andreesen JR. 1999 Selenium-containing xanthine dehydrogenase from Eubacterium barkeri. Eur. J. Biochem. 264, 862 – 871. (doi:10.1046/j.1432-1327.1999.00678.x)

62. Andrianopoulos A, Brons J, Davis MA, Hynes MJ. 2006 TheamdA regulatory gene of Aspergillus nidulans: characterization of gain-of-function mutations and identification of binding sites for the gene product. Fungal Genet. Biol. 21, 50 – 63. (doi:10.1006/fgb.1997.0968)

63. Murphy RL, Andrianopoulos A, Davis MA, Hynes MJ. 1997 Identification of amdA, a new Cys–2–His–2 (C2H2) zinc-finger gene involved in the regulation of the amdS gene of Aspergillus nidulans. Mol. Microbiol. 23, 591 – 602. (doi:10.1046/j.1365-2958.1997.00172.x)

64. Shimanuki M, Uehara L, Pluskal T, Yoshida T, Kukobu A, Kawasaki Y, Yasunaga M. 2013 Klf1, a C2H2 zinc-finger transcription factor, is required for cell wall maintenance during long-term quiescence in differentiating G0 phase. PLoS ONE 8, e78545. (doi:10.1371/journal.pone.0078545)

65. Poch O. 1997 Conservation of a putative inhibitory domain in the GAL4 family members. Gene 184, 229 – 235. (doi:10.1016/S0378-1119(96)00602-6)

66. Stone G, Sadowski I. 1993 GAL4 is regulated by a glucose-responsive functional domain. EMBO J. 12, 1375 – 1385.

67. Gallmeier E, et al. 2015 Reversible oxidation of a conserved methionine in the nuclear export sequence determines subcellular distribution and activity of the fungal nitrate regulator NitA. PLoS Genet. 11, e1005297. (doi:10.1371/journal.pgen.1005297)

68. McGary KL, Slot JC, Rokas A. 2013 Physical linkage of metabolic genes in fungi is an adaptation against the accumulation of toxic intermediate compounds. Proc. Natl Acad. Sci. USA 110, 481 – 486. (doi:10.1073/pnas.1304461110)

69. Scacciozio C. 2014 Fungal biology in the post-genomic era. Fungal Biol. Biotechnol. 1, 7 – 24. (doi:10.1186/s40694-014-0007-6)

70. Wissecaer JH, Slot JC, Rokas A. 2014 The evolution of fungal metabolic pathways. PLoS Genet. 10, e1004816. (doi:10.1371/journal.pgen.1004816)

71. Slot JC, Hibbert DS. 2007 Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. PLoS ONE 2, e1097. (doi:10.1371/journal.pone.0001097)

72. de Oliveira IM, Henriques JA, Bonatto D. 2007 In silico identification of a new group of specific bacterial and fungal nitroreductases-like proteins. Bioch. Biophys. Res. Commun. 355, 919 – 925. (doi:10.1016/j.bbrc.2007.02.049)

73. Roldán MD, Perez-Reinaido E, Castillo F, Moreno-Vivian C. 2008 Reduction of polystyrene compounds: the bacterial nitroreductases. FEMS Microbiol. Rev. 32, 474 – 500. (doi:10.1111/j.1574-6976.2008.01070.x)

74. Flippin M, Sun J, Robellet X, Karaffa L, Fekete E, Zeng AP, Kubicek CP. 2009 Biodiversity and evolution of primary carbon metabolism in Aspergillus nidulans and other Aspergillus spp. Fungal Genet. Biol. 46(Suppl 1), S19 – S59. (doi:10.1016/j.fgb.2008.07.018)

75. Boycheva S, Daviet L, Wolfender JL, Fitzpatrick TB. 2014 The rise of operon-like gene clusters in plants. Trends Plant Sci. 19, 447 – 459. (doi:10.1016/j.tplants.2014.01.013)

76. Mayford ST et al. 2013 Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat. Plant Cell 25, 1078 – 1092. (doi:10.1105/tpc.113.110551)

77. Nutzmann HW, Osbourn A. 2014 Gene clustering in plant specialized metabolism. Curr. Opin. Biotechnol. 26, 91 – 99. (doi:10.1016/j.copbio.2013.10.009)

78. Cove DJ. 1966 The induction and repression of nitrogen reductase in the fungus Aspergillus nidulans. Biochem. Biophys. Acta 173, 51 – 56. (doi:10.1016/S0006-3604(63)80020-0)

79. Bal J, Kajtaniak EMJPN. 1977 4-Nitroquinoline-1-oxide: a good mutagen for Aspergillus nidulans. Micologia 69:838.2008.00107.x)

80. Karacsony Z, Gacser A, Vagvolgyi C, Scaccozio C, Hamari Z. 2014 A dually located multi-HMG-box protein of Aspergillus nidulans has a crucial role in
conidial and ascospore germination. Mol. Microbiol. 94, 383–402. (doi:10.1111/mmi.12727)

89. Bradford MM. 1976 A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254. (doi:10.1016/0003-2697(76)90527-3)

90. Laemmli UK. 1970 Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685. (doi:10.1038/227680a0)

91. Lewis NJ, Scazzocchio C. 1977 The genetic control of molybdenum-containing proteins in Aspergillus nidulans. A xanthine dehydrogenase I half-molecule in cnx-molybdoflavoproteins in Aspergillus nidulans. Eur. J. Biochem. 76, 441–446. (doi:10.1111/j.1432-1033.1977.tb11613.x)

92. Specht CA, D’Russa CC, Novotny CP, Ullrich RC. 1982 A method for extracting high-molecular-weight deoxyribonucleic acid from fungi. Anal. Biochem. 119, 138–163. (doi:10.1016/0003-2697(82)90680-7)

93. Sambrook J, Fritsch EF, Maniatis T. 1989 Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

94. Yanisch-Perron C, Vieira J, Messing J. 1985 Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Nucleic Acids Res. 13, 1033.1977.tb11613.x)

95. Chaveroche MK, Ghigo JM, d’Enfert C. 2000 A rapid method for efficient gene replacement in the filamentous fungus Aspergillus nidulans. Nucleic Acids Res. 28, E97. (doi:10.1093/nar/28.22.e97)

96. Hananah D. 1983 Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol. 166, 557–580. (doi:10.1016/S0022-2836(83)80284-8)

97. Larionov A, Krause A, Miller W. 2005 A standard curve based method for relative real time PCR data processing. BMC Bioinformatics 6, 62. (doi:10.1186/1471-2105-6-62)

98. Yu JH, Han KH, Seo JA, Reyes-Dominguez Y, Scazzocchio C. 2000 Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi. Fungal Genet. Biol. 41, 973–981. (doi:10.1016/j.fgb.2004.08.001)

99. Criscuolo A, Grifaldo S. 2010 BMGE (Block Mapping and Gathering with Entropy): a new software for the selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. 10, 210. (doi:10.1186/1471-2148-10-210)

100. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2008 Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Syst. Biol. 59, 307–321. (doi:10.1093/sysbio/yq010)

101. Dereeper A et al. 2008 Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36, W465–W469. (doi:10.1093/nar/gkn180)

102. Anisimova M, Guindel O. 2006 Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst. Biol. 55, 539–552. (doi:10.1080/10635150600755453)

103. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. 2007 Prediction of subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. 300, 1005–1016. (doi:10.1006/jmbi.2000.3903)

104. Schluter A, Real-Chicharro A, Gabaldon T, Sanchez-Jimenez F, Pujol A. 2010 PeroxisomeDB 2.0: an integrative view of the global peroxisomal metabolome. Nucleic Acids Res. 38, DB00–DB05. (doi:10.1093/nar/gkp593)

105. Kosugi S, Hasebe M, Tomita M, Yanagawa H. 2009 Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proc. Natl Acad. Sci. USA 106, 10 171–10 176. (doi:10.1073/pnas. 090064106)

106. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. 2007 WoLF PSORT: protein localization predictor. Nucleic Acids Res. 35, W585–W587. (doi:10.1093/nar/gkm259)

107. King BR, Guda C. 2007 nqLOC: an n-gram-based Bayesian method for estimating the subcellular proteomes of eukaryotes. Genome Biol. 8, R68. (doi:10.1186/gb-2007-8-5-r68)

108. Guex N, Peitsch MC. 1997 SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18, 2714–2723. (doi:10.1186/gb-2007-8-5-r68)

109. Zhang Y. 2008 I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9, 40. (doi:10.1186/1471-2105-9-40)

110. Roy A, Kucukural A, Zhang Y. 2010 I-TASSER: a unified platform for automated protein structure prediction. Nat. Protoc. 5, 725–738. (doi:10.1038/nprot.2010.5)

111. Humphrey W, Dalke A, Schulten K. 1996 VMD: visual molecular dynamics. J. Mol. Graph. 14, 27 – 38. (doi:10.1016/0267-8894(96)00018-5)

112. Roberts E, Eargle J, Wright D, Luthey-Schulte Z. 2006 MultiSeq: unifying sequence and structure processing. BMC Bioinformatics 7, 62. (doi:10.1186/1471-2105-7-62)

113. King BR, Guda C. 2007 nqLOC: an n-gram-based Bayesian method for estimating the subcellular proteomes of eukaryotes. Genome Biol. 8, R68. (doi:10.1186/gb-2007-8-5-r68)