nature genetics

Birth of a metabolic gene cluster in yeast by adaptive gene relocation

Simon Wong & Kenneth H Wolfe

Although most eukaryotic genomes lack operons, they contain some physical clusters of genes that are related in function despite being unrelated in sequence[1–5]. How these clusters are formed during evolution is unknown. The *DAL* cluster is the largest metabolic gene cluster in yeast and consists of six adjacent genes encoding proteins that enable *Saccharomyces cerevisiae* to use allantoin as a nitrogen source[6]. We show here that the *DAL* cluster was assembled, quite recently in evolutionary terms, through a set of genomic rearrangements that happened almost simultaneously. Six of the eight genes involved in allantoin degradation, which were previously scattered around the genome, became relocated to a single subtelomeric site in an ancestor of *S. cerevisiae* and *Saccharomyces castellii*. These genomic rearrangements coincided with a biochemical reorganization of the purine degradation pathway, which switched to importing allantoin instead of urate. This change eliminated urate oxidase, one of several oxygen-consuming enzymes that were lost by yeasts that can grow vigorously in anaerobic conditions. The *DAL* cluster is located in a domain of modified chromatin involving both H2A.Z histone exchange and Hst1-Sum1–mediated histone deacetylation, and it may be a coadapted gene complex formed by epistatic selection.

There is increasing evidence that the order of genes along chromosomes in many eukaryotes is nonrandom[1]. One of the most notable examples occurs in oat plants, where the genes for at least four unrelated enzymes in a species-specific antimicrobial defense pathway show complete genetic cosegregation[2]. In *S. cerevisiae*, statistically significant physical clustering of essential genes[7] and of genes encoding proteins in the same metabolic pathway[3] has been reported. Neighboring yeast genes tend to be coexpressed[8], and neighbors that are coexpressed tend not to be rearranged during evolution[9]. The physical clustering of genes whose functions are related but that were not formed by tandem gene duplication is particularly intriguing. How do these clusters come together, and over what timescale? And what selective pressure drives their formation? To explore these questions we searched for an example of a metabolic gene cluster that became assembled during recent yeast evolution.

We used the KEGG database of metabolic pathways in *S. cerevisiae* to find groups of genes whose products catalyze successive enzymatic steps, and that are located close together in the genome (with no more than three intervening genes; see **Supplementary Methods** online). This identified two groups of three genes each: the *GAL* cluster[10] and the *DAL* cluster[6]. To estimate the significance of this finding, we used computer simulations in which gene order was randomized, and found that only 92 of 100,000 simulated genomes contained two or more three-gene clusters. The *GAL* genes (*GAL1-GAL10-GAL7*, involved in galactose assimilation) are clustered in the genomes of every yeast species in which they are present[10], which is uninformative with regard to how the cluster originated. The *DAL* cluster, however, seemed to have been formed recently and we therefore selected it for further investigation.

Our KEGG analysis identified only *DAL1*, *DAL2* and *DAL3*—the genes encoding the three catabolic enzymes that break down allantoin into urea and glyoxylate—but the *DAL* cluster also contains genes for allantoin permease (*DAL4*), a probable allantoin racemase (*DCG1*) and an isoform of malate synthase (*DAL7*) that disposes of the glyoxylate and prevents product inhibition of the Dal3 reaction[6] (**Fig. 1**). These six genes are located consecutively in a region of 9.4 kb near the right telomere of *S. cerevisiae* chromosome IX (**Fig. 2**). Together they allow yeast to use allantoin, which is a degradation product of purines, as a nitrogen source. The only allantoin pathway genes that are not located in the cluster are *DAL5* (for allantoate permease), *DUR1,2* (encoding a protein that forms ammonia from urea) and several genes encoding transcription factors[6].

To investigate how the *DAL* cluster originated, we compared the locations of homologous genes in other hemiascomycete yeasts (**Fig. 2**). The cluster is completely conserved, and flanked by the same genes, in the four closest relatives of *S. cerevisiae* (the *Saccharomyces sensu stricto* group). In the *sensu lato* species *S. castellii*, the six genes are also clustered, but there are two differences in the internal arrangement of gene order and the cluster is located in a different part of the genome, though probably still subtelomeric (**Fig. 2** and **Supplementary Fig. 1** online).

Notably, there is no *DAL* cluster in any of the more distantly related hemiascomycete species. Instead, homologs of the six *DAL* genes are found individually at scattered locations in their genomes (**Fig. 2** and **Supplementary Fig. 2** online). A detailed comparison

Department of Genetics, Smurfit Institute, University of Dublin, Trinity College, Dublin 2, Ireland. Correspondence should be addressed to K.H.W. (khwolfe@tcd.ie).

Published online 12 June 2005; doi:10.1038/ng1584

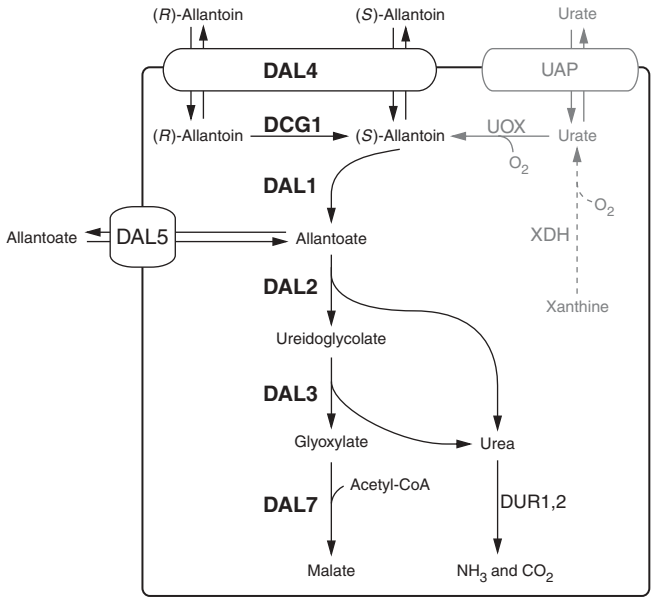

Figure 1 The allantoin degradation pathway. Proteins encoded by genes located in the *DAL* cluster are indicated in bold. *S. cerevisiae* imports allantoin and allantoate, which are degraded to produce ammonia. The pathway of urate import (using urate permease, UAP) and oxidation (using urate oxidase, UOX) that occurs in other yeasts is shown in gray. Xanthine dehydrogenase (*XDH*) genes are present in filamentous ascomycetes such as *Aspergillus nidulans* but not in yeasts. Allantoin exists as two stereoisomers (*R* and *S*), and the assignment of a racemase function to the *DCG1* gene product is based on its sequence similarity to other racemases[6]. The subcellular locations of the *DAL* pathway enzymes in *S. cerevisiae* are not known, but Dal7 has a putative peroxisomal targeting sequence (as does Mls1). In *Candida tropicalis*, the UOX and Dal3 reactions take place in the peroxisome, whereas Dal1 and Dal2 are cytosolic[30].

between *S. cerevisiae* and *Kluyveromyces waltii* (**Fig. 3**) provides an example. Throughout the genome, gene order between *S. cerevisiae* and *K. waltii* is normally well conserved except for the effects of whole-genome duplication in *S. cerevisiae*[11]. Homologs of the genes flanking the *S. cerevisiae DAL* cluster, *YVH1* (*YIR026C*) and *MGA2* (*YIR033W*), are close together in *K. waltii* and are separated only by the chromosome XI gene *RAM2* (*YKL019W*). The *K. waltii YVH1-RAM2-MGA2* region forms part of a large segment of conserved gene order that pairs it with sister regions[11] on *S. cerevisiae* chromosomes IX and XI, but the six *DAL* genes are missing from the expected site and instead are found at six other genomic locations (**Fig. 3**), none of which is near a telomere.

The species that possess a *DAL* cluster form a monophyletic group[10,12] (**Fig. 2**). In contrast, the species without the cluster are paraphyletic and include representatives of at least four lineages (represented by *Yarrowia lipolytica*, *Candida albicans*, *K. waltii* and *Zygosaccharomyces rouxii* in **Fig. 2**) that are independent outgroups to *S. cerevisiae* and *S. castellii*. The most parsimonious interpretation is that the *DAL* cluster was assembled recently, after the split with *Z. rouxii*. Moreover, in species where the *DAL* genes are not clustered, the genes next to each *DAL* gene are often the same (**Fig. 2** and **Supplementary Fig. 2**). For example, *YJR129C* is downstream of *DAL2* in six species. Under the alternative hypothesis that the *DAL* cluster was ancestrally present at the base of the phylogenetic tree and broke up independently several times, one would have to propose that the breakup caused *DAL2* to land beside *YJR129C* on three separate occasions: once on each of the separate lineages represented by *C. albicans*, *K. waltii* and *Z. rouxii*. The conclusion that assembly of a *DAL* cluster is a more parsimonious explanation than disassembly does not depend on knowing the exact branching order among *K. waltii*, *Saccharomyces kluyveri*, *Kluyveromyces lactis* and *Ashbya gossypii*, which is currently unresolved[10,12] and is shown as a polychotomy in **Figure 2**.

Four of the *DAL*-cluster genes (*DAL1*, *DAL2*, *DAL3* and *DCG1*) are single-copy genes in *S. cerevisiae* and seem to have transposed to their new locations in the cluster (they may possibly have been moved by gene duplication followed by loss from the original locus). The other two, *DAL4* and *DAL7*, were produced by duplication of progenitor genes that remain at their original sites in the *S. cerevisiae* genome (**Supplementary Fig. 2**). The allantoin permease gene *DAL4* is a duplicate of the uracil permease gene *FUR4*. In the species lacking a *DAL* cluster, there is one gene orthologous to both *DAL4* and *FUR4*, and its genomic location corresponds to the location of *FUR4* in *S. cerevisiae*. The Dal4 and Fur4 proteins are members of a purine-related transporter family with ten transmembrane helices. They have almost identical lengths and 70% sequence identity, with most of the differences occurring near their N termini. The molecular basis for their different activities is not known. There is no physiologically significant overlap in their functions, but *FUR4* can act as a high-copy suppressor of a *dal4* mutation, which indicates that Fur4 is able to transport allantoin into the cell at a low rate[6]. Similarly, *S. cerevisiae DAL7* is a duplicate of *MLS1* (81% protein sequence identity) and the single *DAL7-MLS1* homolog in species lacking the cluster is found at the equivalent of the *S. cerevisiae MLS1* location. Both *MLS1* and *DAL7* encode malate synthase, but they are regulated very differently (the glyoxylate-cycle gene *MLS1* is glucose repressed whereas *DAL7* is nitrogen repressed[13]). Phylogenetic analysis of both the *DAL4-FUR4* and the *DAL7-MLS1* gene families shows, as expected, that the gene duplications occurred at approximately the same time as the cluster was formed (**Supplementary Fig. 3** online).

The genomic reorganization of the *DAL* genes occurred on a single branch of the phylogenetic tree and coincided with a biochemical reorganization of the purine degradation pathway. There is a phylogenetic correlation between the presence of the cluster and the loss of the genes for urate oxidase (UOX) and urate permease (UAP) in yeast species (**Fig. 2**). In the classical purine degradation pathway, xanthine is converted to urate and then to allantoin, in two successive oxidation steps catalyzed by the peroxisomal enzymes xanthine dehydrogenase (XDH) and UOX (**Fig. 1**). XDH is not present in any yeast species, so to use purine derivatives as a nitrogen source yeasts must import urate, allantoin or allantoate from outside the cell. *K. lactis*, *S. kluyveri* and *Z. rouxii*, which lack the *DAL* cluster, are all able to grow on urate as a sole nitrogen source[14], presumably using UAP to import it, UOX to oxidize it and the *DAL* pathway enzymes to break it down to urea. In contrast, *S. cerevisiae* and *S. castellii* have lost the ability to use urate, have no *UOX* or *UAP* genes, and instead import allantoin using a duplicated (neofunctionalized) copy of *FUR4*. The subsequent degradative steps involve the same *DAL* pathway enzymes in all yeasts, but in *S. cerevisiae* and *S. castellii* the genes have been reorganized into a cluster and *MLS1* has been duplicated to produce *DAL7*.

We suggest that this biochemical reorganization was driven by selection for ability to grow in environments with limited oxygen, which resulted in selection to bypass UOX. The reaction catalyzed by UOX requires molecular oxygen as a substrate (**Fig. 1**) and takes place in the peroxisome[15]. *S. cerevisiae*, the other *sensu stricto* species and *S. castellii* can grow vigorously in anaerobic conditions by fermentation, but the other species listed in **Figure 2** are petite-negative (they cannot

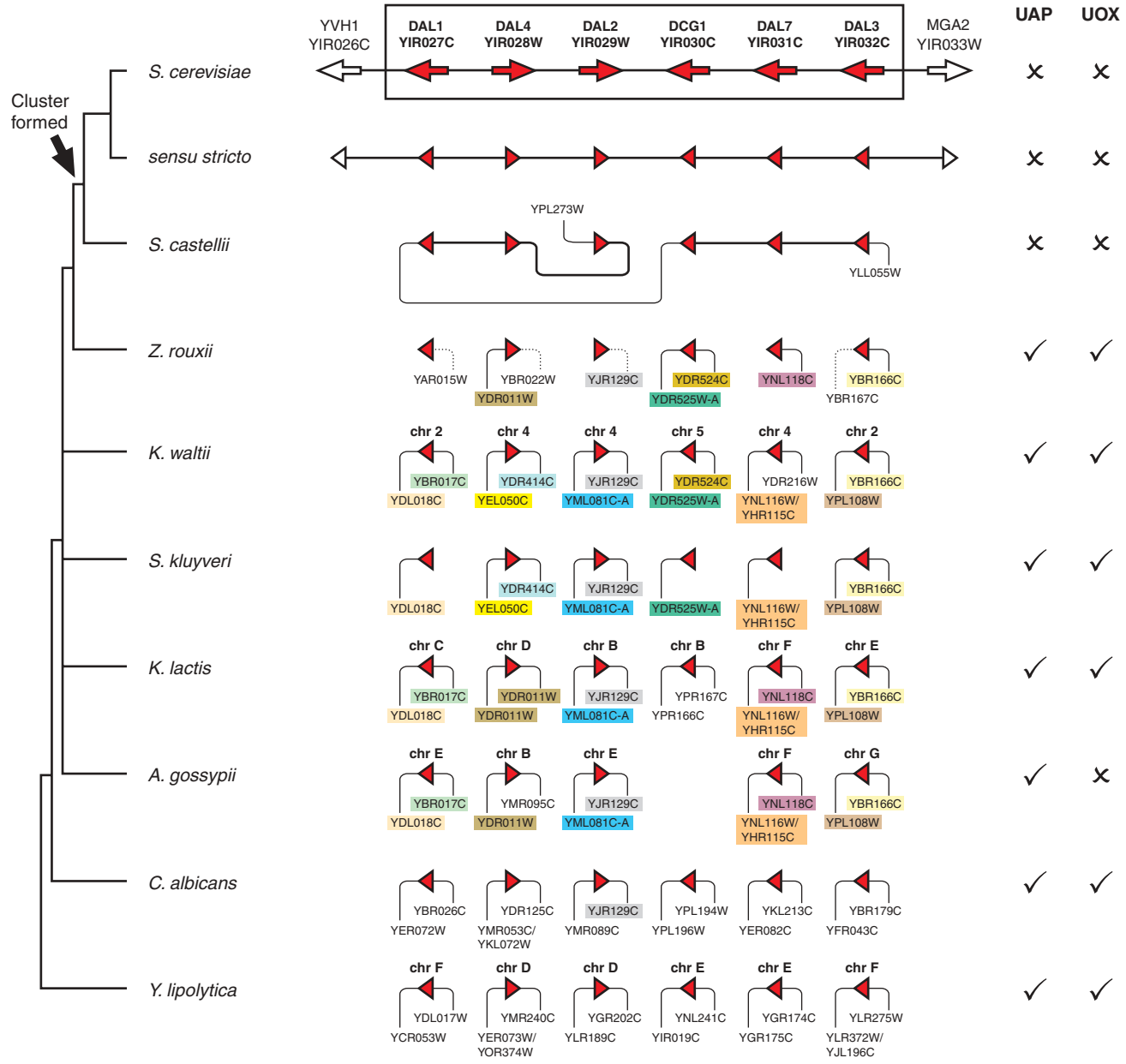

Figure 2 Conservation and disruption of the *DAL* gene cluster. The phylogenetic tree on the left is a conservative consensus of refs. 10,12. The central panel shows the *S. cerevisiae DAL* cluster at the top (boxed) and identifies the genes flanking the orthologs of the *DAL* genes (red triangles) in other species. Color coding is used to show occurrence of the same flanking genes in different species. The *sensu stricto* line indicates conservation of the cluster in *Saccharomyces paradoxus*, *Saccharomyces mikatae*, *Saccharomyces kudriavzevii* and *Saccharomyces bayanus*. The panel on the right indicates the presence or absence of urate oxidase (*UOX*) and urate permease (*UAP*) genes in each species. Information is based on draft or complete genome sequences of all species except *Z. rouxii*, where more limited survey sequencing was done (see **Supplementary Methods** for sources of sequence data). Dashed lines for *Z. rouxii* indicate gene order inference from paired clone-end sequences. All six *DAL*-cluster genes are absent from the genome of *Candida glabrata*. For more details, see **Supplementary Figure 2**.

lose their mitochondrial DNA) and all of them except for *S. kluyveri* are strict aerobes[16]. Yeast species that are facultative anaerobes have a competitive advantage over strict aerobes in oxygen-poor natural environments such as the interiors of decomposing plant material[17]. It has been suggested that selective pressure to economize on oxygen during evolution of the 'fermentative lifestyle' in *S. cerevisiae* led to a decreasing dependence on oxygen-requiring reactions[16]. Consistent with this hypothesis, *S. cerevisiae* and *S. castellii* have lost copies of several other genes that, like *UOX*, encode peroxisomal enzymes that consume molecular oxygen and are present in most yeast species that lack the *DAL* cluster (**Table 1**). The importance to *S. cerevisiae* of growth in low oxygen is also demonstrated by the way its pyrimidine synthesis pathway has been reconfigured to avoid dependence on respiration[16] and by the existence of many duplicated genes encoding specialized hypoxic and aerobic forms of proteins[18]. We therefore propose that selection for less dependence on oxygen led to a switch

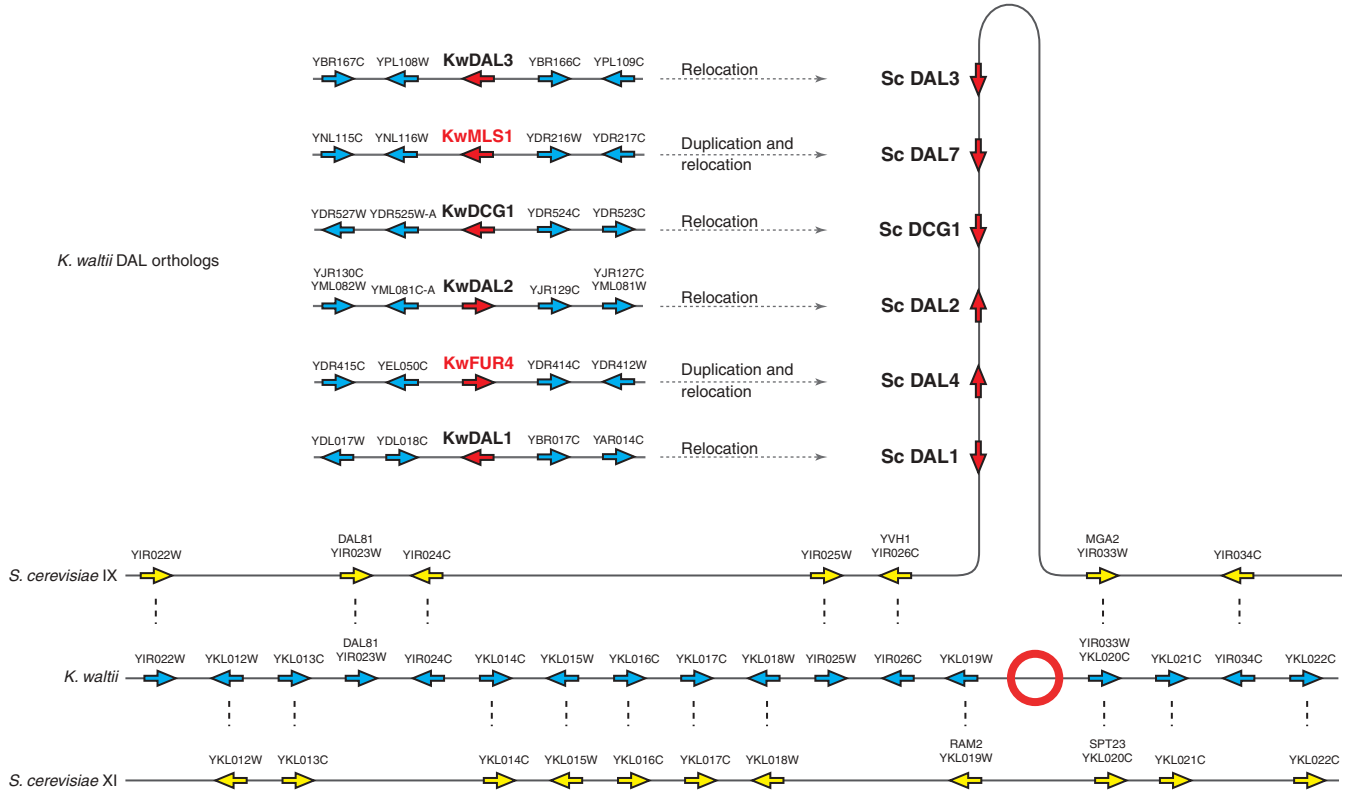

Figure 3 Locations of the *DAL* genes in *K. waltii* as compared with *S. cerevisiae*. The *DAL* cluster on *S. cerevisiae* chromosome IX lies within a sister region[11] between chromosomes IX and XI (genes shown in yellow, including the duplicated gene pair *MGA2* and *SPT23*). The corresponding region in *K. waltii* (blue) contains a merge of genes from the two chromosomes, but the *DAL* genes (red) are missing from the expected site in *K. waltii* (red circle). Instead, orthologs of the each of the six *DAL* genes are found in separate *K. waltii* genomic regions as shown in the upper left part of the figure. *DAL1*, *DAL2*, *DAL3* and *DCG1* apparently transposed to the cluster site, whereas *DAL4* and *DAL7* were formed by duplication of *FUR4* and *MLS1*.

from urate to allantoin utilization in an ancestor of *S. cerevisiae* and *S. castellii*. Natural sources of allantoin for yeasts are plants[19] and insect excretion[20].

Biochemical reorganization of the purine degradation pathway in response to evolutionary pressure to lose UOX can explain why the gene duplication that produced *DAL4* (the allantoin permease gene) was advantageous; but why did the genes in the pathway become relocated into a cluster at the same time? Gene clusters can be formed by selection either for physical proximity of the genes or for genetic linkage[7]. The *DAL* cluster has attributes that suggest that both of these processes may have been involved in its formation, as discussed later.

The *DAL* cluster may have been formed by selection for genetic linkage, to keep particular combinations of *DAL* gene alleles in linkage disequilibrium[7]. If so, the cluster could be regarded as an example of a coadapted gene complex[1]. When the fitness of an allele at one locus depends on the genotype at another locus, this can result in a selective advantage for genomic rearrangements that bring two favorably interacting alleles into tight linkage disequilibrium. This kind of epistatic selection has been shown in the case where each of the alleles alone is deleterious in the absence of the other[21]. Gene clusters that have been formed by epistatic selection are expected to be recombinational cold spots[7], and the meiotic recombination rate in the *DAL* cluster is indeed lower than in 97% of similar-sized regions in the yeast genome (**Supplementary Fig. 4** and **Supplementary Note** online). The reorganization of the purine degradation pathway (**Fig. 1**) involved the creation of two new genes (*DAL4* and *DAL7*) by duplication. The new duplicates, while adapting to their new roles in the allantoin degradation pathway, may have interacted well with only particular alleles at other *DAL* loci. Under epistatic selection, rearrangement of these genes into a cluster would be favored.

Table 1 Numbers of genes in yeast species for peroxisomal enzymes that use molecular oxygen as a substrate

Gene	*S. cer.*	*S. cas.*	*Z. rou.*	*K. wal.*	*S. klu.*	*K. lac.*	*A. gos.*	*C. alb.*	*Y. lip.*
Urate oxidase (UOX)	0	0	1	1	1	1	0	1	1
Copper-containing amine oxidase (CAO)	0	0	1	1	1	1	1	2	2
D-amino acid oxidase (DAO)	0	1	3	2	3	3	0	3	4
Long-chain fatty alcohol oxidase (FAO)	0	0	0	0	0	0	0	1	0
Fatty acyl–CoA oxidase (POX1)	1	1	1	1	1	1	1	3	6

A further factor that may have led to epistatic selection for linkage is the fact that glyoxylate, which is produced by the Dal3 reaction and removed by the Dal7 reaction, contains a reactive aldehyde group and is quite toxic to yeast[6], so there may be selection for alleles of *DAL3* and *DAL7* that interact well and facilitate metabolic channeling[2]. Channeling is suggested by the finding that Dal3 enzymatic activity is reduced in a *dal7* mutant[13]. It is notable that *DAL3* and *DAL7* are immediate neighbors in both *S. cerevisiae* and *S. castellii* (**Fig. 2**).

Selection for physical clustering is suggested by the location of the *DAL* cluster in a chromosomal region that is subject to chromatin modification, as are many of the subtelomeric regions in yeast[22]. Subtelomeric regions also tend to contain genes that permit growth on unusual nutrient sources and whose expression is induced under appropriate stress conditions[22]. Transcription of the *DAL* genes is strongly induced by nitrogen limitation[6]. Yeast cells growing in natural conditions carry substantial reserves of allantoin sequestered in vacuoles, which can be mobilized if no external source of nitrogen is available[6]. The *DAL* cluster is subject to two unusual chromatin modifications. First, the whole cluster (from *YVH1* to *DAL3*; **Fig. 2**) comprises one of the largest of 18 Htz1-activated domains (HZADs) where the normal histone H2A in nucleosomes has been replaced by the variant H2A.Z (Htz1)[23]. This exchange is effected by the Swr1 chromatin remodeling complex[24], and the HZAD forms the boundary between heterochromatin (the region of Sir2-effected silencing closer to the telomere) and euchromatin. Detailed analysis[23,24] of this HZAD shows that H2A.Z is present in (and enables expression of) the genes *YVH1*, *DAL1*, *DAL2*, *DCG1* and *DAL3*, but not *DAL4* or *DAL7*. Second, the intergenic spacer between *DAL1* and *DAL4* binds the transcription repressor Sum1, which in turn specifically recruits Hst1 (a paralog of Sir2), which deacetylates histones H3 and H4. The *DAL1-DAL4* spacer is one of the five strongest sites for Hst1 binding in the *S. cerevisiae* genome[25], and expression of *DAL1* is induced if Sum1 is deleted[26]. The main role of the Hst1-Sum1 system is to repress expression of mid-sporulation genes during mitotic growth, and this repression can be lifted upon starvation for nitrogen, which induces meiosis. We therefore suggest that, in addition to the documented regulation of *DAL* genes by nitrogen catabolite repression[6], the genes in the *DAL* cluster are also coordinately regulated by chromatin modification involving a site of Hst1-Sum1 activity located within an HZAD. The *DAL* region is the only place in the yeast genome where an Hst1 site occurs in an HZAD[23,25]. Assuming that there is a selective advantage to repressing each of the *DAL* genes when nitrogen is not limiting, there would have been an incremental selective advantage to relocating each gene into the chromatin modification domain. The hypotheses of genetic and physical clustering discussed here are not necessarily incompatible; one way in which alleles could interact well is by being amenable to the same type of chromatin modification.

Our findings point to an episode of intense selective pressure that resulted in reorganization of the *DAL* genes into a cluster at one particular point during yeast evolution. In contrast to the adaptive evolution of their location, the DAL genes do not show evidence of positive selection on their sequences (**Supplementary Table 1** online). The pressure to relocate the genes was probably a consequence of the biochemical reorganization of the upstream steps in the purine degradation pathway (**Fig. 1**), which in turn may have been a response to selection for decreased oxygen consumption. It is likely that the subtelomeric location of the *DAL* cluster greatly facilitated its assembly, because subtelomeres have several properties that predispose them to being a workshop for evolutionary genomic experimentation. DNA from other places in the genome is frequently duplicated to subtelomeres[27], and they are often polymorphic within a species with respect to gene presence or absence[28]. They show high rates of gene duplication and interchromosomal DNA exchange during evolution[27]. Despite this dynamism, their meiotic recombination rate is low[29]. These properties facilitate both the formation of new and unusual combinations of sequences and the retention of useful combinations in linkage disequilibrium. Lastly, subtelomeres contain many domains of modified chromatin, which enables groups of genes to be coordinately induced or repressed and allows for adaptation by epigenetic mechanisms. We do not know whether this process is representative of how other gene clusters are formed, but because many of the properties of yeast subtelomeres are shared by heterochromatic regions in other eukaryotes, it is possible that heterochromatin could frequently be involved in the assembly of eukaryotic gene clusters.

METHODS

See **Supplementary Methods** for details of methodology.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank D. Scannell, K. Byrne, J. Gordon, G. Conant, N. Khaldi, B. Cusack, L. Rusche, J. Conery and anonymous referees for comments. This study was supported by Science Foundation Ireland.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 11 November 2004; accepted 2 May 2005
Published online at http://www.nature.com/naturegenetics/

1. Hurst, L.D., Pal, C. & Lercher, M.J. The evolutionary dynamics of eukaryotic gene order. *Nat. Rev. Genet.* **5**, 299–310 (2004).
2. Qi, X. *et al.* A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc. Natl. Acad. Sci. USA* **101**, 8233–8238 (2004).
3. Lee, J.M. & Sonnhammer, E.L. Genomic gene clustering analysis of pathways in eukaryotes. *Genome Res.* **13**, 875–882 (2003).
4. Keller, N.P. & Hohn, T.M. Metabolic pathway gene clusters in filamentous fungi. *Fungal Genet. Biol.* **21**, 17–29 (1997).
5. Blumenthal, T. & Gleason, K.S. *Caenorhabditis elegans* operons: form and function. *Nat. Rev. Genet.* **4**, 112–120 (2003).
6. Cooper, T.G. Regulation of allantoin catabolism in *Saccharomyces cerevisiae*. in *The Mycota III: Biochemistry and Molecular Biology* (ed. Marzluf, G.A.) 139–169 (Springer, Berlin, 1996).
7. Pal, C. & Hurst, L.D. Evidence for co-evolution of gene order and recombination rate. *Nat. Genet.* **33**, 392–395 (2003).
8. Cohen, B.A., Mitra, R.D., Hughes, J.D. & Church, G.M. A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat. Genet.* **26**, 183–186 (2000).
9. Hurst, L.D., Williams, E.J. & Pal, C. Natural selection promotes the conservation of linkage of co-expressed genes. *Trends Genet.* **18**, 604–606 (2002).
10. Hittinger, C.T., Rokas, A. & Carroll, S.B. Parallel inactivation of multiple *GAL* pathway genes and ecological diversification in yeasts. *Proc. Natl. Acad. Sci. USA* **101**, 14144–14149 (2004).
11. Kellis, M., Birren, B.W. & Lander, E.S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**, 617–624 (2004).
12. Kurtzman, C.P. & Robnett, C.J. Phylogenetic relationships among yeasts of the '*Saccharomyces* complex' determined from multigene sequence analyses. *FEMS Yeast Res.* **3**, 417–432 (2003).
13. Hartig, A. *et al.* Differentially regulated malate synthase genes participate in carbon and nitrogen metabolism of *S. cerevisiae*. *Nucleic Acids Res.* **20**, 5677–5686 (1992).
14. LaRue, T.A. & Spencer, J.F. The utilization of purines and pyrimidines by yeasts. *Can. J. Microbiol.* **14**, 79–86 (1968).
15. van der Klei, I.J. & Veenhuis, M. Yeast peroxisomes: function and biogenesis of a versatile cell organelle. *Trends Microbiol.* **5**, 502–509 (1997).
16. Piskur, J. & Langkjaer, R.B. Yeast genome sequencing: the power of comparative genomics. *Mol. Microbiol.* **53**, 381–389 (2004).
17. Deak, T. & Beuchat, L.R. *Handbook of Food Spoilage Yeasts* (CRC Press, Boca Raton, Florida, USA, 1996).
18. Kwast, K.E. *et al.* Genomic analyses of anaerobically induced genes in *Saccharomyces cerevisiae*: functional roles of Rox1 and other factors in mediating the anoxic response. *J. Bacteriol.* **184**, 250–265 (2002).

19. Peoples, M.B. & Gifford, R.M. Regulation of the transport of nitrogen and carbon in higher plants. in *Plant Metabolism* (eds. Dennis, D.T., Layzell, D.B., Lefebvre, D.D. & Turpin, D.H.) 525–538 (Longman, Singapore, 1997).
20. Bursell, E. The excretion of nitrogen in insects. *Adv. Insect Physiol.* **4**, 33–67 (1967).
21. Nei, M. Modification of linkage intensity by natural selection. *Genetics* **57**, 625–641 (1967).
22. Robyr, D. *et al.* Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. *Cell* **109**, 437–446 (2002).
23. Meneghini, M.D., Wu, M. & Madhani, H.D. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* **112**, 725–736 (2003).
24. Krogan, N.J. *et al.* A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol. Cell* **12**, 1565–1576 (2003).
25. Robert, F. *et al.* Global position and recruitment of HATs and HDACs in the yeast genome. *Mol. Cell* **16**, 199–209 (2004).
26. Pierce, M. *et al.* Sum1 and Ndt80 proteins compete for binding to middle sporulation element sequences that control meiotic gene expression. *Mol. Cell. Biol.* **23**, 4814–4825 (2003).
27. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E.S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241–254 (2003).
28. Daran-Lapujade, P. *et al.* Comparative genotyping of the *Saccharomyces cerevisiae* laboratory strains S288C and CEN.PK113–7D using oligonucleotide microarrays. *FEMS Yeast Res.* **4**, 259–269 (2003).
29. Gerton, J.L. *et al.* Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **97**, 11383–11390 (2000).
30. Takada, Y. & Tsukiji, N. Peroxisomal localization and activation by bivalent metal ions of ureidoglycolate lyase, the enzyme involved in urate degradation in *Candida tropicalis*. *J. Bacteriol.* **169**, 2284–2286 (1987).