Minireview

Rewiring the transcriptional regulatory circuits of cells

Devin R Scannell and Ken Wolfe

Address: Department of Genetics, Smurfit Institute, University of Dublin, Trinity College, Dublin 2, Ireland.

Correspondence: Ken Wolfe. E-mail: khwolfe@tcd.ie

Published: 29 January 2004

Genome ***Biology*** 2004, **5:**206

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2004/5/2/206

Abstract

The molecular mechanisms that regulate gene expression can evolve either by changing the *cis*-acting DNA elements in promoters, or by replacing the *trans*-acting regulatory proteins. New data from yeast species show that both processes can happen.

Whether you're driving a Cadillac or an electric golf buggy, you put your foot on the accelerator and it goes. Most of us don't care what's happening in the engine so long as it does what we want it to do, and which vehicle we choose depends on whether we want to drive across the country or across the fairway. Natural selection is a bit like us uninformed motorists: it favors organisms that work well in a particular context, without prejudice as to what is the best mechanism to achieve this. But the catch about evolution is that every organism has been produced by descent with modification from ancestor organisms that worked, so an engineer cannot just design your 'golf buggy' from scratch. To push the analogy, it has to be made from a Model T Ford by continually adding, subtracting or modifying parts, keeping the car running all the time.

If you find it hard to see how an engineer could do this, you will appreciate why we have a limited understanding of how evolution can alter the regulatory circuitry of cells. Despite the facts that a study of gene regulation in the *lac* operon was one of the first great successes of molecular biology [1], and that it has long been argued that the molecular differences between species are more likely to hinge on gene regulatory differences than on gene content differences [2], few comprehensive studies into how gene regulation evolves have been attempted. The reason for this is no secret. Whereas the coding region of the *lac* operon consists of three genes totaling 4,941 nucleotides, the operator and promoter sequences comprise only 160 bases [3]. In addition, *cis*-regulatory elements do not evolve within a clearly defined structural framework such as that imposed by the triplet code, nor do they exhibit the close relationship between sequence and function that exists in coding regions. So, although incredible progress has been made in the last decade in molecular evolutionary studies of genes and genomes (see [4] for a review), how gene regulation evolves is still a topic about which there is more speculation than hard data.

On the other hand, there is more to regulatory evolution than just *cis* elements. In this regard, a recent paper by Tsong *et al.* [5] on the evolution of the yeast mating-type circuitry provides a valuable lesson. By showing that one yeast species uses a regulatory protein that is simply not present in another species, Tsong *et al.* [5] bring home the message that gene regulation is achieved by *cis* elements working in conjunction with their cognate transcription factors, with both sets of factors operating in the specific context of a genetic circuit, the final output of which is the object of natural selection. The system studied by Tsong *et al.* is the *MTL* (mating-type-like) locus of the pathogenic yeast *Candida albicans*, which corresponds to the *MAT* (mating-type) locus of *Saccharomyces cerevisiae* (Figure 1). The *S. cerevisiae* system has been studied intensively and comprises two genes, *MATα1* and *MATα2*, encoded by the α version of the locus, and one gene, *MATa1*, encoded by the **a** version; each haploid cell is either **a** or α type. The combination of genes for one activator protein (α1) and two repressors (α2 and the **a**1/α2 dimer) in diploid cells allows the

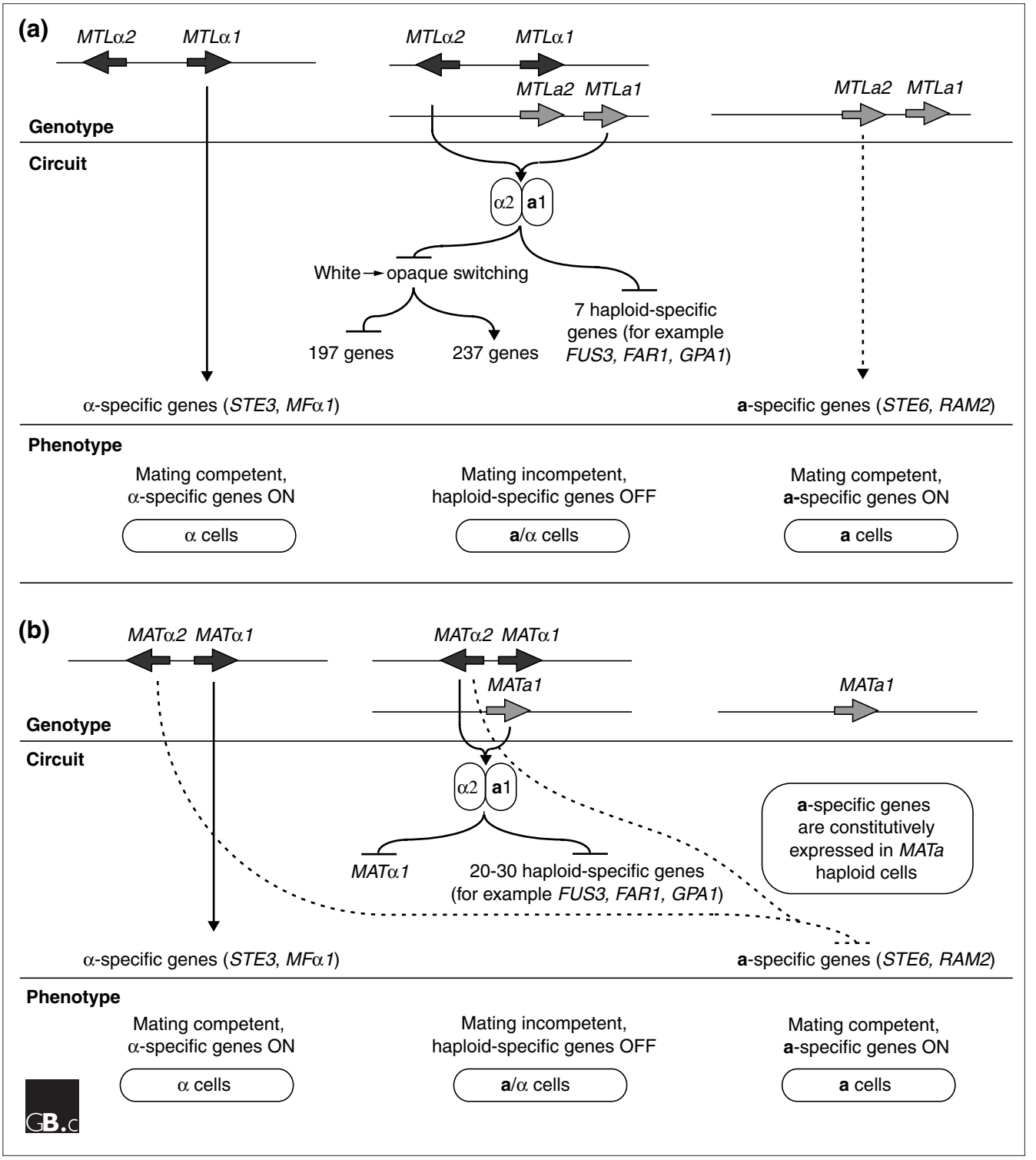

Figure 1
Genotypes, interactions and outputs of the mating-type loci of *Candida albicans* and *Saccharomyces cerevisiae* [5,15]. **(a)** The *C. albicans MTL* locus. The α version of the locus encodes the α1 protein, which activates the expression of α-specific genes; the **a** version of the locus encodes the **a**2 protein, which activates the expression of **a**-specific genes. In **a**/α cells, where both versions of the locus are present, the α2 and **a**1 proteins act together as a repressor of the phenotypic switch from white to opaque, and of 'haploid-specific' genes. **(b)** The *S. cerevisiae MAT* locus. As in *C. albicans*, the α1 protein activates the expression of α-specific genes, and α2 and **a**1 act together in diploid cells as a repressor of *MATα1* and haploid-specific genes. The **a**-specific genes are, however, constitutively expressed in the absence of the α2 repressor. The dotted lines highlight the replacement of a positive branch in *C. albicans* with a negative one in *S. cerevisiae*.

S. cerevisiae MAT locus to produce three types of output, corresponding to α maters, **a** maters, and non-mating **a**/α diploid cell types. In *C. albicans*, however, the **a** idiomorph contains a second gene, *MTLa2*, that is completely absent from *S. cerevisiae*.

At the start of the recent analysis, Tsong *et al.* [5] knew only two things about *MTLa2*. The first is that it encodes a DNA-binding protein with an HMG domain (named for the 'high mobility group' proteins in which it was first found), unrelated in sequence to any of the *S. cerevisiae* mating-type genes. The second is that it is not responsible for the most obvious difference in mating between the two yeasts, namely that in order to mate a *C. albicans* cell must undergo a morphological change from the white phase, which is thought to aid its avoidance of the host's immune system, to the mating-competent opaque phase. *MTLa1* and *MTLα2* together had previously been demonstrated by Miller and Johnson [6] to be necessary and sufficient to regulate (repress) phenotypic switching between white and opaque morphologies. With the exception of this initial requirement, however, mating in the two species was thought to proceed by similar mechanisms.

In order to fully understand the genetic circuit regulating mating, and in particular the role of the novel transcriptional regulator *MTLa2*, all sixteen possible knockout combinations of the *MTLa1*, *MTLa2*, *MTLα1* and *MTLα2* genes were generated in *C. albicans* and their mating ability assayed [5]. These experiments showed that, like in *S. cerevisiae*, α1 is a positive regulator of the ability of α-cells to mate, and **a**1 and α2 together repress mating ability. Additionally, they revealed that, unlike in *S. cerevisiae*, in which **a**-specific genes are constitutively expressed unless the α2 protein is present to repress them, in *C. albicans* **a**-specific genes require the novel **a**2 protein as a positive regulator and are insensitive to α2 as a negative regulator. Interestingly, although these last two observations represent regulatory differences between *S. cerevisiae* and *C. albicans*, because the two changes effectively cancel each other out the circuit performs a logically identical operation in the two species and the output remains three cell types (Figure 1).

When genomic data from other yeasts is used to put these observations in a phylogenetic context [5,7], it is clear that the *C. albicans* circuit is the ancestral one and that in *S. cerevisiae* α2 has acquired a new, negative regulatory function to compensate for the loss of **a**2. This rewiring of *MAT*-locus function occurred in the *Saccharomyces* lineage of yeasts soon after they had acquired the Ho endonuclease that allows haploid cells to switch mating types efficiently [7]. In fact, comparisons between yeasts and filamentous ascomycetes such as *Neurospora* suggest that the *MAT* locus remained in the same chromosomal location (beside the *SLA2* gene) for eons while its gene content underwent several changes [7].

In addition to the evolutionary changes in the repertoire and function of regulatory proteins, we can safely infer that there have been multiple instances of change in the *cis* elements that these proteins recognize. In the **a**-specific genes of *S. cerevisiae* many *cis* elements that were previously acted on by the **a**2 protein must have been lost, and multiple sites for repression by α2 must have been gained. For example, the *RAM2* and *STE6* genes, which encode proteins involved in farnesylation and secretion of **a**-factor pheromone, respectively, have shifted from being activated by **a**2 (as in *C. albicans*) to being repressed by α2 (as in *S. cerevisiae*) [5]. At present we do not know whether this shift occurred by co-opting new *cis* elements, or by retaining the same *cis* elements but changing the proteins that bind them. It will therefore be interesting to determine the DNA-binding specificity of *C. albicans* **a**2.

Tsong *et al.* [5] also used microarrays to dissect the transcriptional consequences of the phenotypic switch from white to opaque. In *S. cerevisiae*, the **a**1/α2 dimer, which is only present in heterozygous diploids, represses 20-30 genes associated with mating. By contrast, in *C. albicans* this dimer represses only seven genes directly but approximately 400 others indirectly through control of the white to opaque morphological switch. The functions of these 400 genes are not restricted to involvement in mating, but also include aspects of the pathogenic lifestyle of *C. albicans*. Again, the phylogenetic context suggests an interpretation: during the evolution of virulence in *C. albicans* (a derived characteristic of this lineage), these 400 genes were brought under the control of a much older genetic circuit, *MTL*. The question of how hundreds of genes could be marshaled by a new regulator is an interesting one. The evidence suggests that it happened as an indirect consequence of changing a small number of direct targets of the **a**1/α2 heterodimer. This kind of parsimony in rewiring genetic circuits has also been observed in artificial evolution experiments in *S. cerevisiae*. Following 500 generations of glucose-limited growth, Ferea *et al.* [8] observed that many genes involved in energy pathways and carbohydrate metabolism were coordinately up- or down-regulated. Correlated changes in the expression profiles of known regulators of these genes support the idea that, rather than selection for multiple independent downstream changes, a few changes further upstream in the regulatory hierarchy are responsible.

The logic of the *MAT* locus circuit discussed above shares many similarities with the genetic circuit underlying wing polyphenism in ants. Polyphenism is the ability of a genome to produce different phenotypes in response to different external signals, as dramatically illustrated by winged queen ants and their wingless worker sisters. In ants, polyphenism is achieved by interrupting, in workers, the genetic circuit that would otherwise execute a program of wing patterning [9]. The circuit is believed to have evolved only once, and is conserved among insect species (including *Drosophila*) that

last shared a common ancestor over 300 million years ago. In spite of this, Abouheif and Wray [9] found that the point at which the program is interrupted is different in different ant species, and is labile over periods as short as 20 million years. As in the *MAT* locus, the output phenotype (two states of 'wingedness'; three states of 'yeastness') has been conserved across species, but the details of how the regulatory circuits in each species achieve the conserved outputs are different.

Where does this leave us? It is becoming clear that in addition to studies aimed at dissecting the evolutionary conservation of *cis* elements [10-12], we need many more studies in the style of Tsong *et al*. [5] - detailed analyses of real genetic circuits over appropriate evolutionary distances, where both *cis*- and *trans*-acting factors can change (see also Hinman *et al*. [13]). The level of detail is also going to be very important. For instance, a recent analysis of the *lac* operon shows that, contrary to what has been believed for the last 50 years, it has not evolved to permit a smooth transition from fully repressed to fully expressed; instead, the circuit encodes four distinct plateaus and thresholds [14]. It will be vital to have such accurate information in order to be able to determine whether or not a genetic circuit is truly conserved, and to peer a little more closely into the transcriptional engine of the cell.

References

1. Jacob F, Monod J: **Genetic regulatory mechanisms in the synthesis of proteins.** *J Mol Biol* 1961, **3:**318-356.
2. Wilson AC, Carlson SS, White TJ: **Biochemical evolution.** *Annu Rev Biochem* 1977, **46:**573-639.
3. Salgado H, Santos-Zavaleta A, Gama-Castro S, Millan-Zarate D, Diaz-Peredo E, Sanchez-Solano F, Perez-Rueda E, Bonavides-Martinez C, Collado-Vides J: **RegulonDB (version 3.2): transcriptional regulation and operon organization in *Escherichia coli* K-12.** *Nucleic Acids Res* 2001, **29:**72-74.
4. Wolfe KH, Li W-H: **Molecular evolution meets the genomics revolution.** *Nat Genet* 2003, **33 Suppl:**255-265.
5. Tsong AE, Miller MG, Raisner RM, Johnson AD: **Evolution of a combinatorial transcriptional circuit: a case study in yeasts.** *Cell* 2003, **115:**389-399.
6. Miller MG, Johnson AD: **White-opaque switching in *Candida albicans* is controlled by mating-type locus homeodomain proteins and allows efficient mating.** *Cell* 2002, **110:**293-302.
7. Butler G, Kenny C, Fagan A, Kurischko C, Gaillardin C, Wolfe KH: **Evolution of the *MAT* locus and its Ho endonuclease in yeast species.** *Proc Natl Acad Sci USA*, DOI: 10.1073/pnas.0303793101.
8. Ferea TL, Botstein D, Brown PO, Rosenzweig RF: **Systematic changes in gene expression patterns following adaptive evolution in yeast.** *Proc Natl Acad Sci USA* 1999, **96:**9721-9726.
9. Abouheif E, Wray GA: **Evolution of the gene network underlying wing polyphenism in ants.** *Science* 2002, **297:**249-252.
10. Cliften P, Sudarsanam P, Desikan A, Fulton L, Fulton B, Majors J, Waterston R, Cohen BA, Johnston M: **Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting.** *Science* 2003, **301:**71-76.
11. Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES: **Sequencing and comparison of yeast species to identify genes and regulatory elements.** *Nature* 2003, **423:**241-254.
12. Chiang DY, Moses AM, Kellis M, Lander ES, Eisen MB: **Phylogenetically and spatially conserved word pairs associated with gene-expression changes in yeasts.** *Genome Biol* 2003, **4:**R43.
13. Hinman VF, Nguyen AT, Cameron RA, Davidson EH: **Developmental gene regulatory network architecture across 500 million years of echinoderm evolution.** *Proc Natl Acad Sci USA* 2003, **100:**13356-13361.
14. Setty Y, Mayo AE, Surette MG, Alon U: **Detailed map of a *cis*-regulatory input function.** *Proc Natl Acad Sci USA* 2003, **100:**7702-7707.
15. Johnson AD: **Molecular mechanisms of cell-type determination in budding yeast.** *Curr Opin Genet Dev* 1995, **5:**552-558.