Following Gene Duplication, Paralog Interference Constrains Transcriptional Circuit Evolution

Christopher R. Baker *et al.*

Science **342**, 104 (2013);
DOI: 10.1126/science.1240810

been far less. This is qualitatively consistent with conditions necessary to maintain a temperature of ~15°C at Earth's surface with a mixture of CO_2 and other greenhouse gases (*28*, *29*).

References and Notes

1. C. Goldblatt *et al*., *Nat. Geosci.* **2**, 891–896 (2009).
2. R. Wordsworth, R. Pierrehumbert, *Science* **339**, 64–67 (2013).
3. H. I. M. Lichtenegger *et al*., *Icarus* **210**, 1–7 (2010).
4. V. Busigny, P. Cartigny, P. Philippot, *Geochim. Cosmochim. Acta* **75**, 7502–7521 (2011).
5. R. O. Pepin, *Earth Planet. Sci. Lett.* **252**, 1–14 (2006).
6. The excellent match between the atmospheric $^{38}Ar/^{36}Ar$ ratio (two Ar isotopes of primordial origin) and that of Ar trapped in primitive meteorites attests that atmospheric Ar was not quantitatively lost to space [noble gas isotopic ratios are sensitive to thermal and nonthermal escape processes (*5*)]. Independently, noble gas isotope systematics indicate that mantle degassing did not contribute quantitatively to the atmospheric inventory of these elements after Earth's formation. The large difference in the $^{40}Ar/^{36}Ar$ values between the mantle and the atmosphere implies that less than 2% of mantle ^{36}Ar has been transferred to the atmosphere since the time of Earth's formation (*31*). A similar conclusion is obtained from ^{36}Ar flux estimates. The modern flux of ^{36}Ar from the mantle is ~2000 mol/year [from scaling to the present-day ^{3}He flux from the mantle (*32*)]. Integrated over 3.5 Gy (the age of the formations studied here) and assuming a constant flux through time, the amount of ^{36}Ar is only about 0.1% of the ^{36}Ar atmospheric inventory (5.55×10^{15} mol). The degassing rate of the mantle was probably higher in the past because of a higher thermal regime. Coltice *et al*. (*33*) estimated that the melting rate of the mantle was a factor of ~20 higher 3.5 Gy ago as compared to the present-day rate. In this case, the mantle contribution of degassed mantle ^{36}Ar to the atmosphere represents only 0.8% of the ^{36}Ar inventory, which is still negligible in the context of this study.
7. R. Buick, J. S. R. Dunlop, *Sedimentology* **37**, 247–277 (1990).
8. M. J. Van Kranendonk, P. Philippot, K. Lepot, S. Bodorkos, F. Pirajno, *Precambrian Res.* **167**, 93–124 (2008).
9. J. Foriel *et al*., *Earth Planet. Sci. Lett.* **228**, 451–463 (2004).
10. P. Philippot *et al*., *C. R. Palevol* **8**, 649–663 (2009).
11. N. Thébaud, P. Philippot, P. Rey, J. Cauzid, *Contrib. Mineral. Petrol.* **152**, 485–503 (2006).
12. N. Thébaud *et al*., *Earth Planet. Sci. Lett.* **272**, 639–655 (2008).
13. M. Nishizawa, Y. Sano, Y. Ueno, S. Maruyama, *Earth Planet. Sci. Lett.* **254**, 332–344 (2007).
14. D. L. Pinti, K. Hashizume, J. I. Matsuda, *Geochim. Cosmochim. Acta* **65**, 2301–2315 (2001).
15. M. Pujol, B. Marty, R. Burgess, *Earth Planet. Sci. Lett.* **308**, 298–306 (2011).
16. M. Pujol, B. Marty, R. Burgess, G. Turner, P. Philippot, *Nature* **498**, 87–90 (2013).
17. P. G. Burnard, R. Hu, G. Turner, X. W. Bi, *Geochim. Cosmochim. Acta* **63**, 1595–1604 (1999).
18. R. C. Hamme, R. S. Emerson, *Deep Sea Res.* **51**, 1517–1528 (2004).
19. F. Robert, M. Chaussidon, *Nature* **443**, 969–972 (2006).
20. Supplementary materials on *Science* Online.
21. M. A. Kendrick, R. Burgess, R. A. D. Pattrick, G. Turner, *Chem. Geol.* **177**, 351–370 (2001).
22. B. Marty, N. Dauphas, *Earth Planet. Sci. Lett.* **206**, 397–410 (2003).
23. Y. Sano, C. T. Pillinger, *Geochem. J.* **24**, 315–325 (1990).
24. Y. F. Jia, R. Kerrich, *Terra Nova* **16**, 102–108 (2004).
25. T. Owen *et al*., *J. Geophys. Res.* **82**, 4635–4639 (1977).
26. H. B. Niemann *et al*., *Nature* **438**, 779–784 (2005).
27. J. A. Tarduno, R. D. Cottrell, M. K. Watkeys, D. Bauch, *Nature* **446**, 657–660 (2007).
28. C. Sagan, G. Mullen, *Science* **177**, 52–56 (1972).
29. J. F. Kasting, *Nature* **464**, 687–689 (2010).
30. S. M. Som, D. C. Catling, J. P. Harnmeijer, P. M. Polivka, R. Buick, *Nature* **484**, 359–362 (2012).
31. M. Ozima, F. A. Podosek, *Noble Gas Geochemistry* (Cambridge Univ. Press, Cambridge, 2002), and references therein.
32. B. Marty, P. Allé, in *Noble Gas Geochemistry and Cosmochemistry*, J. Matsuda, Ed. (Terra Scientific, Tokyo, 1994), pp. 191–204.
33. N. Coltice, B. Marty, R. Yokochi, *Chem. Geol.* **266**, 4–9 (2009).

Acknowledgments: This study was supported by Agence Nationale de la Recherche Grant eLife-2 to P.P. and B.M. and by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 grant agreement no. 267255 to B.M.). P.P. acknowledges the Institut de Physique du Globe de Paris, the Institut des Sciences de l'Univers, and the Geological Survey of Western Australia for supporting the Pilbara Drilling Project. We thank six anonymous reviewers and D. Catling for constructive comments. This is Centre de Recherches Pétrographiques et Géochimiques contribution no. 2248 and Institut de Physique du Globe de Paris contribution no. 3424.

Supplementary Materials

www.sciencemag.org/content/342/6154/101/suppl/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S3
References (*34*–*41*)

23 May 2013; accepted 5 September 2013
Published online 19 September 2013;
10.1126/science.1240971

Following Gene Duplication, Paralog Interference Constrains Transcriptional Circuit Evolution

Christopher R. Baker,[1,2] Victor Hanson-Smith,[1,2] Alexander D. Johnson[1,2]*

Most models of gene duplication assume that the ancestral functions of the preduplication gene are independent and can therefore be neatly partitioned between descendant paralogs. However, many gene products, such as transcriptional regulators, are components within cooperative assemblies; here, we show that a natural consequence of duplication and divergence of such proteins can be competitive interference between the paralogs. Our example is based on the duplication of the essential MADS-box transcriptional regulator Mcm1, which is found in all fungi and regulates a large set of genes. We show that a set of historical amino acid sequence substitutions minimized paralog interference in contemporary species and, in doing so, increased the molecular complexity of this gene regulatory network. We propose that paralog interference is a common constraint on gene duplicate evolution, and its resolution, which can generate additional regulatory complexity, is needed to stabilize duplicated genes in the genome.

Gene duplications are an important source of new genes, and a variety of models have been developed to rationalize why certain gene duplicates have been maintained over evolutionary time (*1*–*3*). For instance, the neofunctionalization model posits that soon after duplication, one of the duplicates evolves a new function that can be selected for and, thereby, maintained over time (*2*, *3*). Alternatively, subfunctionalization (via the duplication-degeneration-complementation model) holds that duplicates can be maintained in the genome by acquiring reciprocal loss-of-function mutations, such that both duplicates become necessary to perform the combined functions of the preduplication ancestor (*1*–*3*). Classically, these models have assumed that ancestral functions can be treated independently, making the partitioning of these functions among the descendant paralogs possible without detrimental effects (*2*). However, for the many gene products that participate in cooperative assemblies, the molecular interactions that underlie gene functions are not intrinsically independent (*4*). For example, many transcriptional regulators depend on a cooperative network of protein-protein and protein–nucleic acid interactions. In these instances, loss of one or more ancestral molecular interactions will often give rise to competitive interference between gene duplicates (paralog interference) (*5*). Although in some instances this competition may be advantageous, we suspect that paralog interference following gene duplication would typically have detrimental effects that must be evolutionarily bypassed for the paralogs to be maintained. Because many proteins form cooperative assemblies, resolution of paralog interference is likely to be a widespread phenomenon influencing the fate of duplicated genes.

Mcm1 is a fungal MADS-box transcriptional regulator that binds DNA cooperatively with seven different partner transcriptional regulators (cofactors) to control the expression of many genes, including those coding for mating functions and arginine metabolic enzymes (*6*). The way in which

[1]Department of Immunology and Microbiology, University of California, San Francisco, CA 94143, USA. [2]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158, USA.

*Corresponding author. E-mail: ajohnson@cgl.ucsf.edu

Mcm1 assembles at the arginine metabolism (*ARG*) genes varies between fungal clades. In the yeasts *Kluyveromyces lactis* and *Candida albicans*, an Mcm1 homodimer regulates transcription of *ARG* genes by binding specifically to DNA with the cofactor Arg81 (Fig. 1A) (*7*, *8*). In the lineage leading to baker's yeast (*Saccharomyces cerevisiae*), a tandem gene duplication event introduced an extra copy of Mcm1 (called Arg80), such that the

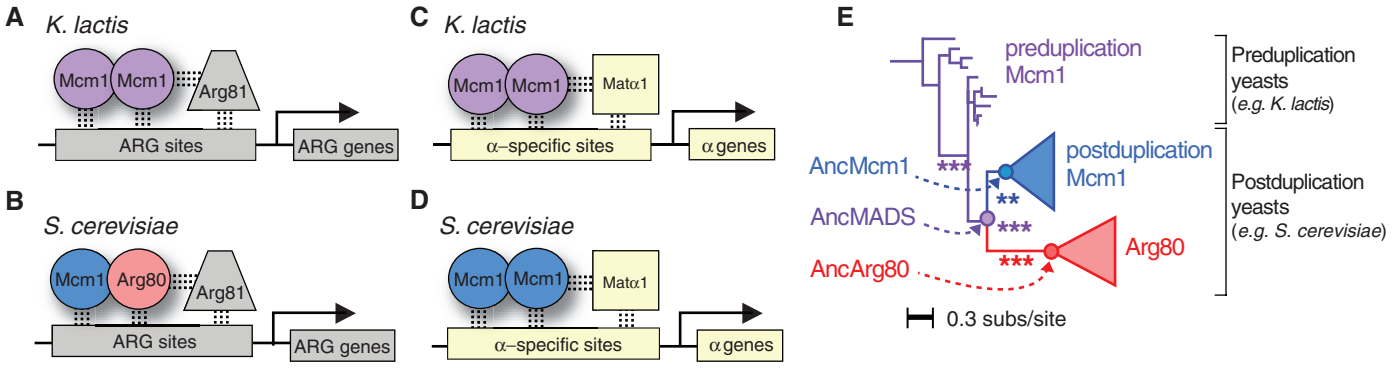

Fig. 1. Function and evolution of MADS-box proteins in hemiascomycete yeasts. (**A**) In *K. lactis*, an Mcm1 homodimer regulates the *ARG* genes by interacting with Arg81 and binding a specific DNA sequence. (**B**) In *S. cerevisiae*, an Mcm1-Arg80 heterodimer interacts with Arg81 to regulate *ARG* genes. (**C**) An Mcm1 homodimer interacts with Matα1 to regulate α-specific genes in *K. lactis* and (**D**) *S. cerevisiae*. (**E**) A maximum likelihood phylogeny of MADS-box domain proteins in hemiascomycete yeasts. A tandem gene duplication generated paralogs Mcm1 and Arg80 in the last shared common ancestor of *Zygosaccharomyces rouxii* and *S. cerevisiae*. Circles denote ancestral proteins reconstructed in this study. Asterisks on internal branches correspond to approximate-likelihood ratio support for the monophyly of the descendant clade: *** denotes support > 10.0; ** denotes support > 5.0. subs, substitutions.

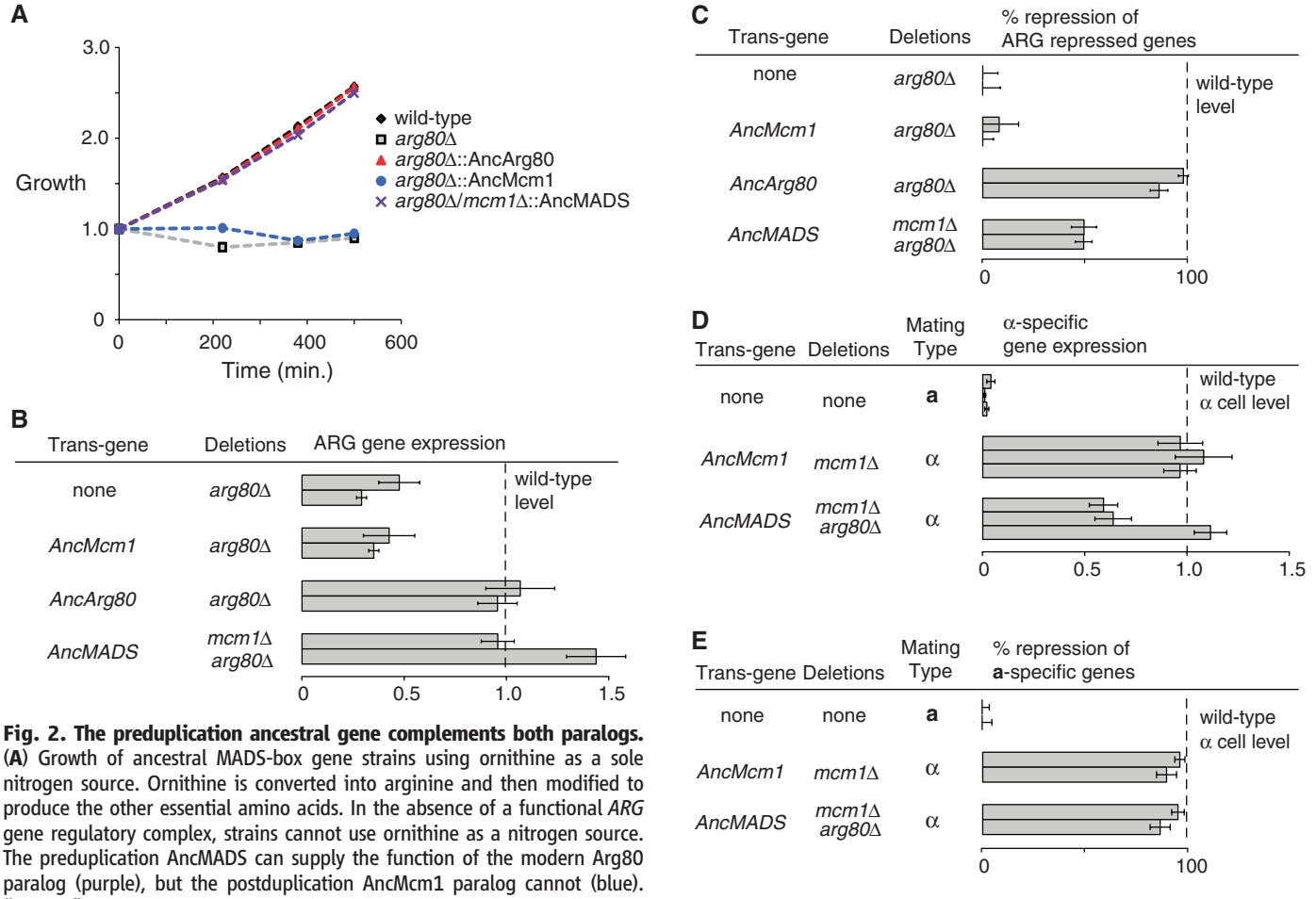

Fig. 2. The preduplication ancestral gene complements both paralogs. (**A**) Growth of ancestral MADS-box gene strains using ornithine as a sole nitrogen source. Ornithine is converted into arginine and then modified to produce the other essential amino acids. In the absence of a functional *ARG* gene regulatory complex, strains cannot use ornithine as a nitrogen source. The preduplication AncMADS can supply the function of the modern Arg80 paralog (purple), but the postduplication AncMcm1 paralog cannot (blue). "Growth" on the *y* axis is the ratio of optical density at 600 nm (OD_{600}) at the indicated time point divided by OD_{600} at time zero. (**B** to **E**) Gene expression profiling of ancestral MADS-box proteins in *S. cerevisiae* quantified with NanoString (www.nanostring.com). (B) MADS-box activated ARG genes. Row 1, *CAR1*; row 2, *CAR2*. (C) MADS-box repressed ARG genes. Row 1, *ARG3*; row 2, *ARG5,6*. (D) MADS-box activated mating genes (α-specific genes). Row 1, *SAG1*; row 2, *MFa1*; row 3, *STE3*. (E) MADS-box repressed mating genes (**a**-specific genes). Row 1, *STE2*; row 2, *STE6*. In each experiment, mean and standard error (indicated by error bars) were determined using three replicates.

S. cerevisiae regulatory architecture is more complex. In *S. cerevisiae*, an Mcm1-Arg80 heterodimer regulates the transcription of *ARG* genes by binding DNA with the cofactor Arg81 (Fig. 1B) (*9*). Other Mcm1-regulated gene sets in *S. cerevisiae* did not experience an increase in regulatory complexity following gene duplication. For instance, the α-specific genes (genes that give α mating cells their specialized properties) are regulated by an Mcm1 homodimer that binds specifically to DNA with the cofactor Matα1 in species that branch before and after the gene duplication event (Fig. 1, C and D) (*10*–*12*). In all instances, gene regulation by Mcm1 and Arg80 depends on the formation of strong interactions with both cofactors and DNA.

To understand how the linked biochemical functions of DNA and cofactor binding diverged after Mcm1 duplicated, we reconstructed ancestral MADS-box proteins, characterized these ancestral proteins in vivo and in vitro, and identified the mutations through which their functions diversified [see supplementary materials and methods and (*13*)]. Specifically, we reconstructed the MADS-box domains of the most recent common shared ancestor of all postduplication Mcm1 paralogs (AncMcm1); all postduplication Arg80 paralogs (AncArg80); and the preduplication, most recent shared common ancestor of all Mcm1 and Arg80 paralogs (AncMADS) (Fig. 1E, fig. S1, and tables S1 and S2).

We integrated the reconstructed ancestral MADS-box proteins into *S. cerevisiae* and removed the modern copies of *ARG80* and *MCM1* to determine if the ancestral proteins could complement their deletion. Deletion of *S. cerevisiae MCM1* is lethal, and deletion of *S. cerevisiae ARG80* produces defects in arginine metabolism (*14*, *15*). We found that the preduplication AncMADS protein complemented both defects: Replacement of Mcm1 or Arg80 with AncMADS had no impact on growth in either rich media or media with a precursor of arginine as a sole nitrogen source, a phenotype that depends on normal *ARG* gene regulation (Fig. 2A and fig. S2A). We measured the expression levels of a representative set of Mcm1 and Arg80/Mcm1 regulated genes and found that AncMADS restored activation and repression of these genes, and most genes showed the same dynamic range as in the wild type (Fig. 2, B to E, and fig. S2B). The exceptions were a diminished dynamic range of gene expression for the *ARG* repressed genes, a mildly diminished dynamic range of gene expression for the α-specific genes, and stronger activation than the wild type for the ARG activated gene *CAR2* (Fig. 2, B to D). In contrast, the postduplication MADS-box proteins (AncArg80 and AncMcm1) failed to complement deletions of the sister paralogs. Specifically, AncMcm1 did not complement the deletion of the native *ARG80* and, similarly, the presence of AncArg80 alone did not rescue the deletion of *MCM1* (an essential gene) (Fig. 2, A to C, and fig. S2C). The capacity of the preduplication ancestral MADS-box protein to complement the functions of both daughter genes in a modern species, combined with the inability of the postduplication ancestors to do the same, shows that AncMcm1 and AncArg80 acquired degenerative mutations that necessitated the retention of both paralogs over evolutionary time.

We next determined the mutations that underlie the diversification of AncMcm1 and AncArg80 following the duplication of AncMADS. The cofactors Matα1 and Arg81 both interact with the same portion of the MADS-box domain, which we refer to as the cofactor binding pocket (*16*, *17*). We modeled the reconstructed protein sequences for AncMADS, AncMcm1, and AncArg80 onto the structure of *S. cerevisiae* Mcm1 in complex with DNA (*18*) and then compared the sequences within the cofactor binding pocket (fig. S3, A to C). On the lineage from AncMADS to AncMcm1, one substitution occurred in the pocket [Tyr^{33}→Phe^{33} (Y33F); Y, Tyr; F, Phe], and it has been conserved in all Mcm1 descendants (Fig. 3A and table S1). On the lineage from AncMADS to AncArg80, three sequence substitutions occurred within the cofactor binding pocket (T41A, Q42N, F62L; T, Thr; A, Ala, Q, Gln; N, Asn; L, Leu), and each has been strongly conserved in postduplication Arg80 descendant sequences (Fig. 3A and table S1). Previous work has shown that these residues play a critical role in stabilizing the interactions between Arg81 and Arg80, as well as Matα1 and Mcm1 (*16*, *17*). To assess the impact of these changes on the preduplication ancestor, we introduced these mutations into AncMADS and observed

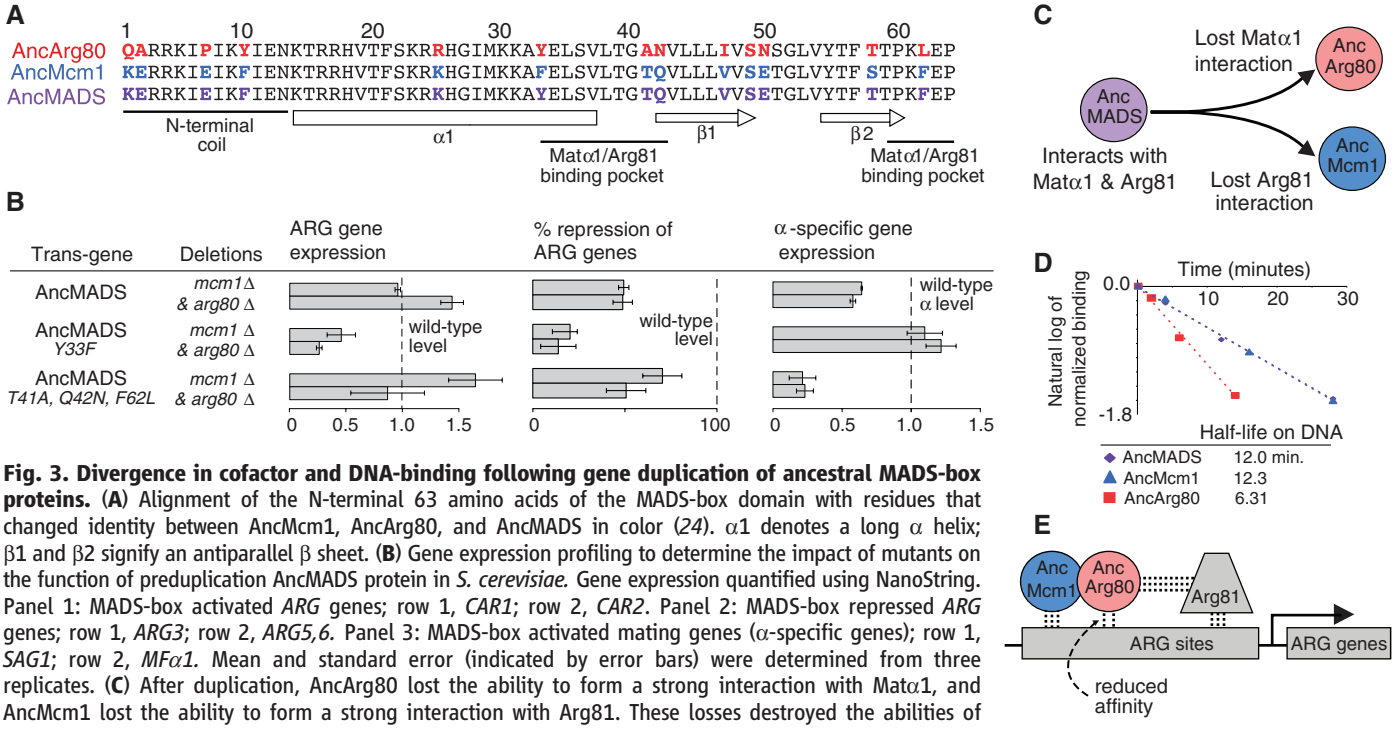

Fig. 3. Divergence in cofactor and DNA-binding following gene duplication of ancestral MADS-box proteins. (**A**) Alignment of the N-terminal 63 amino acids of the MADS-box domain with residues that changed identity between AncMcm1, AncArg80, and AncMADS in color (*24*). α1 denotes a long α helix; β1 and β2 signify an antiparallel β sheet. (**B**) Gene expression profiling to determine the impact of mutants on the function of preduplication AncMADS protein in *S. cerevisiae*. Gene expression quantified using NanoString. Panel 1: MADS-box activated *ARG* genes; row 1, *CAR1*; row 2, *CAR2*. Panel 2: MADS-box repressed *ARG* genes; row 1, *ARG3*; row 2, *ARG5,6*. Panel 3: MADS-box activated mating genes (α-specific genes); row 1, *SAG1*; row 2, *MFα1*. Mean and standard error (indicated by error bars) were determined from three replicates. (**C**) After duplication, AncArg80 lost the ability to form a strong interaction with Matα1, and AncMcm1 lost the ability to form a strong interaction with Arg81. These losses destroyed the abilities of AncArg80 and AncMcm1 to regulate α-specific genes and *ARG* genes, respectively. (**D**) Half-lives of MADS-box ancestors on the *S. cerevisiae CAR2* cis-regulatory sequence (see supplementary materials and methods). Saturating levels of unlabeled DNA were added at time point zero. (**E**) After the duplication of AncMADS, α-specific genes are regulated by a homodimer of AncMcm1, whereas *ARG* genes are regulated by a heterodimer of AncMcm1 and AncArg80 due to the reduced affinity of AncArg80 for DNA.

their effect on expression levels of *ARG* genes and α-specific genes. When we introduced the Mcm1 substitution (Y33F) into AncMADS, we observed a disruption of *ARG* gene regulation but no decrease in α-specific gene regulation; in fact, the dynamic range of α-specific gene regulation slightly expanded (Fig. 3B). When we introduced the Arg80 substitutions (T41A, Q42N, F62L) into AncMADS, we observed decreased α-specific gene regulation but no compromise in *ARG* gene regulation (Fig. 3B). (We note that *CAR2* even returns to wild-type expression levels from its elevated state in the AncMADS background.) We observed similar effects of these mutations on AncMADS when we swapped in the cofactors Arg81 and Matα1 from the preduplication species *K. lactis*, indicating that changes to these two cofactors did not play a major role in partitioning the ancestral molecular interactions of AncMADS (fig. S3, D and E). Taken together, these results show that the preduplication AncMADS could form cofactor interactions with both Arg81 and Matα1 and that these interactions reciprocally degenerated in the descendant paralogs: AncMcm1 lost its ability to productively interact with Arg81, whereas AncArg80 lost its ability to interact with Matα1 (Fig. 3C).

We next examined the DNA-binding surfaces of the pre- and postduplication MADS-box proteins. *S. cerevisiae* Arg80 and Mcm1 have very closely related DNA-binding specificities (*19*, *20*), indicating that, at most, a limited divergence in MADS-box DNA-binding specificity occurred following duplication. That the DNA-binding specificity did not change substantially is also supported by our observation that the preduplication AncMADS protein can complement the deletion of either postduplication gene (Fig. 2). However, *S. cerevisiae* Arg80 has a substantially lower affinity for DNA than Mcm1 (*15*, *16*). To determine when this affinity change occurred, we compared the DNA-binding affinities of AncMADS, AncArg80, and AncMcm1 by measuring their half-lives on DNA. To make the comparisons meaningful, we used an endogenous *S. cerevisiae* MADS-box binding site (taken from an arginine metabolic gene) that closely resembles the consensus site for *S. cerevisiae* Mcm1 and Arg80 and that has been shown to support binding by both these proteins in vitro (*8*, *20*). We observed that the half-life of AncArg80 was significantly lower than the half-lives of AncMADS and AncMcm1 (Fig. 3D). Thus, the difference in affinity between modern Mcm1 and Arg80 is due to a decrease in Arg80 DNA-binding affinity that occurred soon after the duplication (Fig. 3E).

Next, we investigated the consequences of this reduction in Arg80 DNA-binding affinity on gene expression. We hypothesized that a version of Arg80 with full DNA-binding strength might interfere with Mcm1 by binding to α-specific gene regulatory sites and acting as a dominant negative mutant by preventing Mcm1 from binding cooperatively with Matα1 (Fig. 4A). If this were true, then the reduction of Arg80 DNA-binding affinity would minimize competition by weighting DNA-binding at the α-specific genes in favor of Mcm1. To test this idea, we increased the DNA-binding affinity of AncArg80 to that of the preduplication protein and measured the extent of competitive interference with Mcm1 in *S. cerevisiae*. We identified a total of five mutations (K1Q, E2A, E7P, F10Y, K25R; K, Lys; E, Glu; P, Pro; R, Arg) that occurred on the DNA-binding surface of AncArg80 after the AncMADS duplication (Fig. 3A). A subset of these residues are known to affect MADS-box DNA-binding affinity in *S. cerevisiae* (*14*, *16*). We reversed these mutations in AncArg80, returning the DNA-binding region of AncArg80 to its preduplication, high-affinity state, and then measured α-specific gene expression in an *S. cerevisiae* strain lacking the native Arg80. As predicted by our hypothesis (paralog interference), the AncArg80 mutant significantly reduced α-specific gene expression (Fig. 4B). When overexpressed, the nonmutant, low-affinity AncArg80 protein dampened α-specific gene expression, and the overexpressed, high-affinity AncArg80 mutant blocked the expression of the α-specific genes almost entirely (Fig. 4B). [These effects were not an indirect consequence of altering *MCM1* gene expression levels (fig. S4A).] The antagonism between Arg80 and Mcm1 persists in an attenuated form in contemporary species, as the deletion of *ARG80* in *S. cerevisiae* slightly increases α-specific gene expression (Fig. 4B and fig. S4B). On the basis of these observations, we conclude that the historical reduction in Arg80 DNA-binding affinity limited the degree of paralog interference between Arg80 and Mcm1 at the α-specific genes.

The diminished DNA-binding affinity of Arg80 also provides a simple explanation for the origins of the Arg80-Mcm1 heterodimer (as opposed to an Arg80 homodimer) at the ARG genes in *S. cerevisiae*. The cofactor Arg81 contacts only a single (proximal) subunit within the MADS-box dimer, and this interaction will favor Arg80 because of its strong interaction with Arg81. The energetics will favor Mcm1, rather than Arg80, as the second (distal) subunit because of its higher affinity for DNA (Fig. 3E).

By combining ancestral gene reconstructions with the biochemical and genetic tools available for yeasts, we have shown that competition between paralogs arose as an intrinsic consequence of the duplication and subfunctionalization of a deeply conserved transcriptional regulator. This interference was minimized by a set of historical amino acid substitutions, and we suggest that this was necessary for both paralogs to be maintained, as without the weakened affinity of Arg80 for DNA, gene regulation is severely compromised (Fig. 4B). The minimization of interference was accompanied by an increase in regulatory complexity: The increased number of distinct subunits needed to regulate the *ARG* genes in *S. cerevisiae* relative to the preduplication ancestor (three versus two) is necessary to compensate for the reduced DNA-binding affinity of Arg80. Although we do not know whether the mutations that affected protein-protein interactions occurred before, after, or in concert with those that affected DNA-binding, we have shown that each mutation is a loss-of-function (or at least a reduction-of-function) amino acid substitution (*21*, *22*).

For the many gene products that form cooperative assemblies, exemplified by the proteins studied here, ancestral functions depend on a network of molecular interactions. Following gene duplication, the loss of ancestral interactions by such proteins, resulting in subfunctionalization, may unavoidably give rise to dominant negative

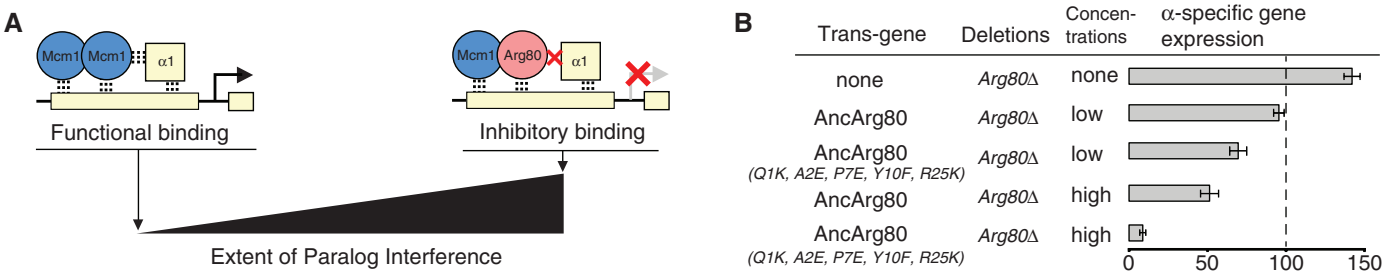

Fig. 4. Paralog interference between Arg80 and Mcm1. (**A**) The ratio of functional binding to inhibitory binding determines the extent of competitive interference between Mcm1 and Arg80. (**B**) Arg80, AncArg80, and AncArg80 mutant with DNA-binding surface changes reverted to ancestral state interfere with α-specific gene expression. α-specific gene expression was quantified by quantitative reverse-transcriptase fluorescence polymerase chain reaction using *SAG1* transcript (normalized to *URA6* transcript). Endogenous expression is driven by the native *ARG80* promoter, and overexpression is driven by the *TEF1* promoter. For gene expression experiments, mean and standard error (indicated by error bars) were determined from five replicates.

effects between duplicates. Although in some cases, this interference can be exploited, for example, by using it to repress gene expression (*5*, *23*), we propose that a more common outcome is the minimization of this interference in gene duplicates that persist over evolutionary time. Whether such minimization is generally accompanied by an increase in regulatory complexity, as seen here, remains to be determined.

References and Notes

1. A. Force *et al.*, *Genetics* **151**, 1531–1545 (1999).
2. H. Innan, F. Kondrashov, *Nat. Rev. Genet.* **11**, 97–108 (2010).
3. A. Wagner, *Genome Biol.* **3**, reviews1012.1–reviews1012.3 (2002).
4. A. Wagner, *Proc. Biol. Sci.* **270**, 457–466 (2003).
5. J. T. Bridgham, J. E. Brown, A. Rodríguez-Marí, J. M. Catchen, J. W. Thornton, *PLOS Genet.* **4**, e1000191 (2008).
6. F. Messenguy, E. Dubois, *Gene* **316**, 1–21 (2003).
7. C. Boonchird, F. Messenguy, E. Dubois, *Mol. Gen. Genet.* **226**, 154–166 (1991).
8. B. B. Tuch, D. J. Galgoczy, A. D. Hernday, H. Li, A. D. Johnson, *PLOS Biol.* **6**, e38 (2008).
9. F. Messenguy, E. Dubois, *Mol. Cell. Biol.* **13**, 2586–2592 (1993).
10. A. Bender, G. F. Sprague Jr., *Cell* **50**, 681–691 (1987).
11. A. E. Tsong, M. G. Miller, R. M. Raisner, A. D. Johnson, *Cell* **115**, 389–399 (2003).
12. C. R. Baker, B. B. Tuch, A. D. Johnson, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7493–7498 (2011).
13. M. J. Harms, J. W. Thornton, *Curr. Opin. Struct. Biol.* **20**, 360–366 (2010).
14. T. B. Acton, J. Mead, A. M. Steiner, A. K. Vershon, *Mol. Cell. Biol.* **20**, 1–11 (2000).
15. N. Amar, F. Messenguy, M. El Bakkoury, E. Dubois, *Mol. Cell. Biol.* **20**, 2087–2097 (2000).
16. A. Jamai, E. Dubois, A. K. Vershon, F. Messenguy, *Mol. Cell. Biol.* **22**, 5741–5752 (2002).
17. J. Mead *et al.*, *Mol. Cell. Biol.* **22**, 4607–4621 (2002).
18. S. Tan, T. J. Richmond, *Nature* **391**, 660–666 (1998).
19. T. E. Hayes, P. Sengupta, B. H. Cochran, *Genes Dev.* **2**, 1713–1722 (1988).
20. F. Messenguy, E. Dubois, C. Boonchird, *Mol. Cell. Biol.* **11**, 2852–2863 (1991).
21. G. C. Finnigan, V. Hanson-Smith, T. H. Stevens, J. W. Thornton, *Nature* **481**, 360–364 (2012).
22. M. Lynch, J. S. Conery, *Science* **302**, 1401–1404 (2003).
23. M. Y. Dennis *et al.*, *Cell* **149**, 912–922 (2012).

24. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Acknowledgments: We thank B. Tuch for assistance and comments; O. Homann for assistance on experimental techniques; and C. Nobile, C. Dalal, S. Devika, R. Bennett, L. Mera, and T. Sorrells for comments. This work was funded by grant RO1 GM057049 from the NIH.

Supplementary Materials

www.sciencemag.org/content/342/6154/104/suppl/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S5
References (*25*–*38*)

21 May 2013; accepted 6 September 2013
10.1126/science.1240810

Surviving in a Marine Desert: The Sponge Loop Retains Resources Within Coral Reefs

Jasper M. de Goeij,[1]* Dick van Oevelen,[2] Mark J. A. Vermeij,[3] Ronald Osinga,[4] Jack J. Middelburg,[5] Anton F. P. M. de Goeij,[6] Wim Admiraal[1]

Ever since Darwin's early descriptions of coral reefs, scientists have debated how one of the world's most productive and diverse ecosystems can thrive in the marine equivalent of a desert. It is an enigma how the flux of dissolved organic matter (DOM), the largest resource produced on reefs, is transferred to higher trophic levels. Here we show that sponges make DOM available to fauna by rapidly expelling filter cells as detritus that is subsequently consumed by reef fauna. This "sponge loop" was confirmed in aquarium and in situ food web experiments, using ^{13}C- and ^{15}N-enriched DOM. The DOM-sponge-fauna pathway explains why biological hot spots such as coral reefs persist in oligotrophic seas—the reef's paradox—and has implications for reef ecosystem functioning and conservation strategies.

Coral reefs thrive in oligotrophic tropical seas, but nevertheless belong to the most productive ecosystems on Earth (*1*–*3*). Efficient retention and recycling of carbon and nutrients causes the net production of reefs to be close to zero, despite high gross primary production (*4*). Reef primary producers such as corals and algae release up to 50% of their fixed carbon (*5*, *6*), of which up to 80% immediately dissolves in seawater (*7*). This shunt into the dissolved organic matter (DOM) pool represents a major flow of energy and nutrients on coral reefs (*7*). In the open ocean, microbes enable the transfer of DOM to higher trophic levels through the well-established microbial loop (*8*). Studies on coral reefs have therefore also initially focused on microbes in reef waters and adjacent permeable sediments to understand the fate of DOM in these systems (*7*, *9*–*11*). However, uptake rates by bacterioplankton, in the sense of the microbial loop, are largely insufficient to explain the observed DOM removal on Caribbean and Indo-Pacific reefs (*12*). It therefore remains unclear how the largest source of energy and nutrients on reefs is transferred to higher trophic levels.

Cryptic habitats, for example, the coral reef's crevices and cavities, are identified as major sinks of DOM on Caribbean and Indo-Pacific reefs (*12*). These habitats cover up to two-thirds of the reef's volume, and the biomass of cryptic organisms can exceed that on the open reef (*13*, *14*). DOM removal rates in cryptic habitats on Caribbean reefs (*12*) are comparable to the average gross primary production rates of the entire coral reef ecosystem (*2*). DOM removal rates on Indo-Pacific reefs are lower (*12*) but still account for up to 46% of the average gross reef productivity. Sponges are primarily responsible for total DOM uptake and remove the same amount of DOM from the water column in 30 min as free-living bacteria take up in 30 days (*12*, *15*). Therefore, sponges retain organic matter within the reef community and thereby prevent energy and nutrient losses to the open ocean. Surprisingly however, sponges respire only 42% of the carbon taken up from the surrounding water (*15*, *16*). Assuming that the remaining 58% is used for growth, a biomass increase of 38% of body carbon per day (more than a doubling of biomass every 3 days) would be expected (*16*). In reality, however, the net growth rate of sponges is near zero (*15*, *16*), implying high losses of sponge biomass through a rapid tissue turnover.

A rapid turnover and extensive loss of sponge cells to the surrounding water has been shown for the sponge *Halisarca caerulea (17)*. The sponge's filter cells (choanocytes) divide every 5 to 6 hours, representing the fastest cell cycle found in any multicellular organism to date (*17*). This rapid cell production is counterbalanced by massive shedding of old choanocytes as particulate organic matter (POM or detritus) into the water column (*17*). Massive shedding of POM is also observed in other tropical sponges (*18*, *19*).

[1]Department of Aquatic Ecology and Ecotoxicology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Post Office Box 94248, 1090 GE Amsterdam, Netherlands. [2]Department of Ecosystem Studies, NIOZ Royal Netherlands Institute for Sea Research, NL-4400 AC Yerseke, Netherlands. [3]Department of Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Carmabi, Piscaderabaai z/n Willemstad, Curaçao. [4]Department of Aquaculture and Fisheries, Wageningen University, Post Office Box 338, 6700 AH Wageningen, Netherlands. [5]Faculty of Geosciences, Utrecht University, Budapestlaan 4, 3584 CD Utrecht, Netherlands. [6]Faculty of Health, Medicine and Life Sciences, Maastricht University, Post Office Box 616, 6200 MD Maastricht, Netherlands.

*Corresponding author. E-mail: j.m.degoeij@uva.nl