

forms of GPRA, and their putative downstream signaling molecules may define a new pathway critically altered in asthma. GPRA encodes isoforms that are produced in distinct patterns by bronchial epithelial cells and smooth muscle cells in asthmatic and healthy individuals. In addition, it is expressed by gut epithelia and keratinocytes of the skin, suggesting a potential role in a wider spectrum of allergic diseases.

References and Notes

1. T. Illig, M. Wjst, *Paediatr. Respir. Rev.* **3**, 47 (2002).
2. S. E. Daniels *et al.*, *Nature* **383**, 247 (1996).
3. T. Laitinen *et al.*, *Nat. Genet.* **28**, 87 (2001).

4. N. I. Leaves, S. Bhattacharyya, S. Wiltshire, W. O. Cookson, *Eur. J. Hum. Genet.* **10**, 177 (2002).
5. P. Höglund *et al.*, *Am. J. Hum. Genet.* **57**, 95 (1995).
6. S. Ranta *et al.*, *Nat. Genet.* **23**, 233 (1999).
7. R. Norio, *Hum. Genet.* **112**, 441 (2003).
8. D. E. Reich, E. S. Lander, *Trends Genet.* **17**, 502 (2001).
9. K. Asumalahti *et al.*, *Hum. Mol. Genet.* **11**, 589 (2002).
10. Materials and methods are available as supporting material on *Science Online*.
11. H. T. Toivonen *et al.*, *Am. J. Hum. Genet.* **67**, 133 (2000).
12. V. Pulkkinen *et al.*, unpublished observations.
13. P. Van Erdevegh *et al.*, *Nature* **418**, 426 (2002).
14. Y. Zhang *et al.*, *Nat. Genet.* **34**, 181 (2003).
15. M. Allen *et al.*, *Nat. Genet.* **35**, 258 (2003).
16. We thank M. Daly, M. Dixon, P. Kauppi, S. Knaappila, R. Lehtinen, H. Turunen, and the staff of GeneOS Limited and Finnish Genome Center for their expertise and help. This study has been supported by

Academy of Finland, Finnish National Technology Agency Tekes, Finnish Antituberculosis Association Foundation, Sigrid Jusélius Foundation, and GeneOS Limited. The French Canadian asthma research program was supported by Canadian Institutes of Health Research and the Burroughs Wellcome Fund. J.K. is a member of Biocentrum Helsinki and Center of Excellence for Disease Genetics at University of Helsinki. L.A.L. and J.K. hold equity in GeneOS Limited and are on the board of directors.

Supporting Online Material

www.sciencemag.org/cgi/content/full/304/5668/300/DC1

Materials and Methods
Tables S1 to S6

4 August 2003; accepted 9 March 2004

The *Ashbya gossypii* Genome as a Tool for Mapping the Ancient *Saccharomyces cerevisiae* Genome

Fred S. Dietrich,^{1,2} Sylvia Voegeli,¹ Sophie Brachat,¹
Anita Lerch,¹ Krista Gates,³ Sabine Steiner,¹ Christine Mohr,¹
Rainer Pöhlmann,¹ Philippe Luedi,^{1,2} Sangdun Choi,⁴
Rod A. Wing,⁴ Albert Flavier,³ Thomas D. Gaffney,³
Peter Philippsen^{1*}

We have sequenced and annotated the genome of the filamentous ascomycete *Ashbya gossypii*. With a size of only 9.2 megabases, encoding 4718 protein-coding genes, it is the smallest genome of a free-living eukaryote yet characterized. More than 90% of *A. gossypii* genes show both homology and a particular pattern of synteny with *Saccharomyces cerevisiae*. Analysis of this pattern revealed 300 inversions and translocations that have occurred since divergence of these two species. It also provided compelling evidence that the evolution of *S. cerevisiae* included a whole genome duplication or fusion of two related species and showed, through inferred ancient gene orders, which of the duplicated genes lost one copy and which retained both copies.

The filamentous fungus *Ashbya gossypii* is currently used in industry for the production of vitamin B₂ (1). It is also an attractive model to study filamentous growth, because of its small genome, haploid nuclei, efficient gene targeting, propagation of plasmids, and growth on defined media (2–8). The *A. gossypii* genome project was initiated when conservation of gene order and orientation (synteny) to *Saccharomyces cerevisiae* was noted (9). We wanted to determine the complete gene repertoire for future work with this fungus, and we aimed at using the gene order information to fully explain the origin of gene

cluster duplications in the *S. cerevisiae* genome that were proposed to represent relics of a whole genome doubling followed by extensive genome rearrangements (10, 11).

Details on the sequencing of the *A. gossypii* genome (GenBank accession numbers AE016814 through AE016821) and annotation are available in the supporting online material. The seven chromosomes encode 4718 proteins, 199 tRNA genes, and at least 49 small nuclear RNA (snRNA) genes. The ribosomal DNA carries 40 copies of ribosomal RNA genes sequenced previously (12). The genome lacks transposons and subtelomeric gene repeats, and gene duplications are rare (table S6). The number of protein-coding genes is similar to the 4824 genes found in *Schizosaccharomyces pombe* (13), suggesting that this may be close to the minimum number of genes needed by a free-living fungus. The genome is extremely compact with an average distance between open reading frames (ORFs) of only 341 base pairs, contributing to an average protein-coding gene

size of only 1.9 kb, clearly less than the 2.1-kb average gene size found in *S. cerevisiae* (14), the 2.5 kb found in *S. pombe* (13), and the 3.7 kb found in *Neurospora crassa* (15). The presence of only 221 introns in the entire *A. gossypii* genome, many at identical positions in *S. cerevisiae* homologs, contributes to the compact nature of this genome.

A. gossypii and *S. cerevisiae* diverged more than 100 million years ago, and their genomes differ substantially in GC content (52% for *A. gossypii* and 38% for *S. cerevisiae*). Still, for 95% of the protein-coding sequences of *A. gossypii*, we found homologs in the *S. cerevisiae* genome, the majority (4281 ORFs) at syntenic locations. Only 175 *A. gossypii* protein-coding genes showed homology but not synteny with *S. cerevisiae* genes, and 262 lack homology (table S3). Several genes with no homologs in *S. cerevisiae* have homologs in *S. pombe* (table S4), supporting the idea that they are real genes not or no longer present in *S. cerevisiae*. The annotation of the *A. gossypii* genome also identified gene functions present in *S. pombe* and *S. cerevisiae* but not in *A. gossypii* (table S5). Protein sequence conservation between syntenic homologs of *A. gossypii* and *S. cer-*

Table 1. Centromere assignments based on synteny of the genes flanking the seven *A. gossypii* and 16 *S. cerevisiae* centromeres. Roman numerals indicate the chromosome number in the respective organism. The two remaining centromeres (X and XII) have synteny with regions on chromosomes I and III, with the double break in synteny coming at the expected position of the centromere.

<i>A. gossypii</i> chromosomes	<i>S. cerevisiae</i> chromosomes
I	III, XIV
II	VIII, XI
III	XIII, XV
IV	V, IX
V	II, IV
VI	I, VII
VII	VI, XVI
Noncentromeric Region (I and III)	X, XII

¹Biozentrum der Universität Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland. ²Department of Molecular Genetics and Microbiology, Duke University Medical Center, Research Drive, Durham, NC 27710, USA. ³Syngenta Biotechnology Inc., 3054 Cornwallis Road, Research Triangle Park, NC 27709, USA. ⁴Departments of Agronomy and Biological Sciences, Clemson University, Clemson, SC 29634, USA.

*To whom correspondence should be addressed. E-mail: Peter.philippsen@unibas.ch

evisiae varies considerably, ranging from less than 20% amino acid identity to nearly 100%. In cases where the sequence identity was less than 30%, synteny was particularly useful in the identification of highly diverged orthologs. The marked fluctuation of sequence conservation of syntenic homologs across the entire genome shows that no one region of the genome is more conserved between these species and that the plasticity of the primary protein sequence varies significantly between genes.

More than 90% of the *A. gossypii* genome could be divided into several hundred synteny groups. In each of these groups, the homology relation of single genes or subgroups of genes alternated between two *S. cerevisiae* regions and *A. gossypii*. We refer to this pattern as double synteny. A relatively simple example of double synteny is shown in Fig. 1A. In this group, thirty-three consecutive protein-coding genes of *A. gossypii* chromosome I, flanked by two tRNA genes, align with regions of consecutive genes from *S. cerevisiae* chromosomes XV and XVI. Homologous ORFs typically are of conserved length (table S3).

When both *S. cerevisiae* regions are combined (with the inversion in chromosome XVI reverted), the resulting gene order matches that of *A. gossypii*, including gene orientations, which justifies annotation of these ORFs as

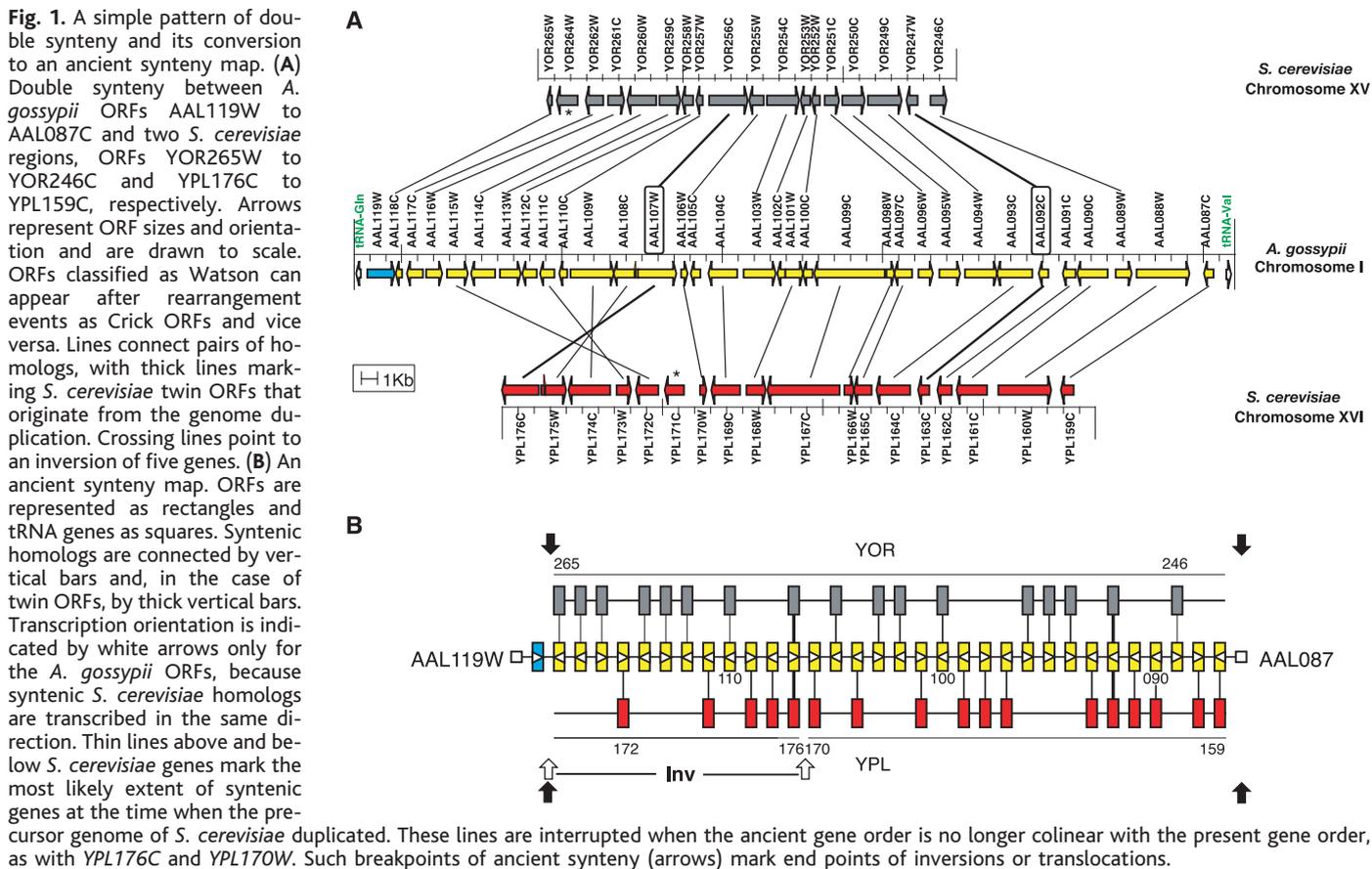
syntenic homologs. This principle has been applied to all ORFs annotated as syntenic homologs in table S3. Three ORFs did not participate in this synteny relation. The *A. gossypii* gene *AAL119W* is a nonsyntenic homolog of *S. cerevisiae* *YFR021W*, a gene involved in vacuolar protein processing. The *S. cerevisiae* gene *YOR264W* plays a role in daughter cell-specific gene expression, a dispensable function in a filamentous fungus, and this may explain the absence of a homolog in *A. gossypii*. *YPL171C* is a nonsyntenic homolog of *AGR329C* and encodes one member of the dehydrogenase family. Nonsyntenic homologs may represent former syntenic homologs, adjacent sequences of which were involved in genome rearrangements so that insufficient evidence of synteny remained.

The pattern of double synteny reflects the gene order of the most recent common ancestor and changes in this order due to genome rearrangements in both lineages, in particular the loss of many genes in the *S. cerevisiae* lineage after the genome duplication. This makes reconstructions of ancient *S. cerevisiae* gene orders relatively easy (Fig. 1B). In this ancient synteny map, protein-coding genes are presented as rectangles and syntenic homologs are connected by vertical bars. All gaps emerging between *S. cerevisiae* genes in this type of presentation indicate positions at the time of the genome

duplication of former homologs that were subsequently lost.

The *A. gossypii* gene order also allowed us to find inversions that occurred in the *S. cerevisiae* lineage, such as the one on chromosome XVI (Fig. 1B), because the alignment of homologs reveals their order before inversion. This ancient gene order shows a single break of synteny, between *YPL176C* and *YPL170W*. The other breakpoint of this inversion coincides with a double break of synteny (black arrows): Two other *S. cerevisiae* gene regions share homology and synteny with the eight *A. gossypii* ORFs proximal to *AAL119W*. The majority of such double breaks of synteny do not coincide with single breaks, and most of them mark end points of inversions or translocations in the evolutionary past of *A. gossypii* or of *S. cerevisiae* before the genome duplication. The *A. gossypii* genome often carries tRNA genes (open squares) or nonsyntenic homologs (blue rectangles) at such end points of rearrangements, as seen in Fig. 1. The presence of tRNA genes at such sites may be explained by the absence of other interspersed repeated DNA elements in *A. gossypii*, and therefore the tRNA genes served as sites for homolog-induced rearrangements.

An example of ancient gene order reconstructed from a complex double synteny pattern is shown in Fig. 2 for the centromere region of



A. gossypii chromosome I. The centromere-proximal genes display homology and synteny to the centromere regions of *S. cerevisiae* chromosomes III (region YCL to YCR) and XIV (region YNL to YNR). A duplication of these two regions was proposed 10 years ago (16) and extensively analyzed on completion of the *S. cerevisiae* genome (10). Other genes in the *A. gossypii* centromere I region show homology and synteny to two short regions of *S. cerevisiae* chromosome XII (YLR) and to a very short region of *S. cerevisiae* chromosome VII (YGL).

Again, all gaps between *S. cerevisiae* genes mark sites of former homologs at the time of genome duplication that were lost during evolution. The alignment of the *S. cerevisiae* homologs revealed eight synteny breaks when compared with today's *S. cerevisiae* genome (14). These breaks originate from three reciprocal translocations (Tra) and one inversion (Inv) in *S. cerevisiae* after genome duplication (white arrows) and from one inversion (black arrows) in the *A. gossypii* lineage (or the *S.*

cerevisiae lineage before genome duplication) (fig. S4).

We aligned the *A. gossypii* and *S. cerevisiae* genomes according to these principles and highlighted synteny breaks. An example of a map of synteny breaks using *A. gossypii* chromosome I as template is given in Fig. 3, and the entire map is given as fig. S5. Close to 96% of the *A. gossypii* genome aligns with two regions from the *S. cerevisiae* genome, including seven duplicated centromere regions and relics of a centromere region down-

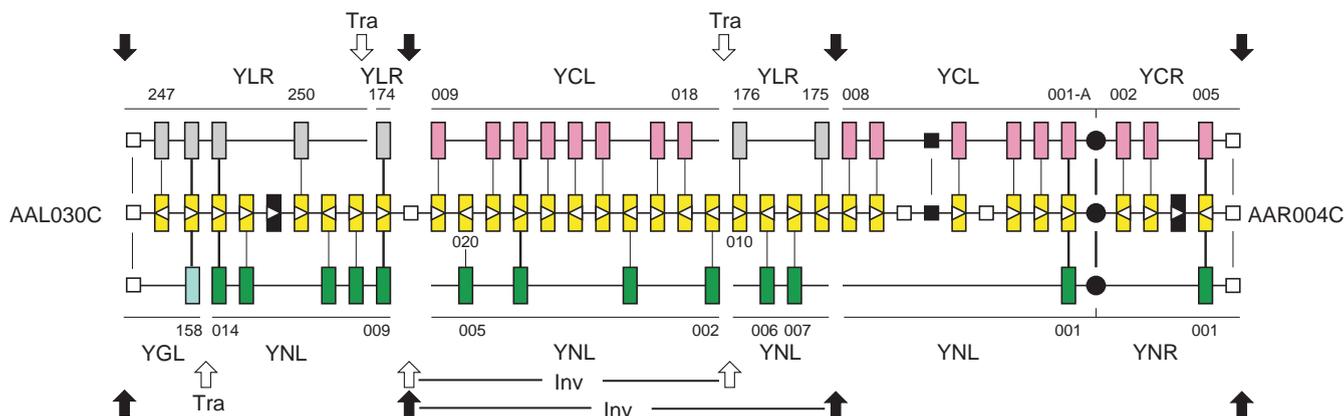


Fig. 2. Multiple rearrangements in an ancient centromere region. *A. gossypii* ORFs AAL030C to AAR004C are connected by vertical lines with those *S. cerevisiae* ORFs that encode homologous proteins. All ORFs participating in this pattern are classified in the GenBank entries as syntenic homologs. They originate from *S. cerevisiae* chromosomes III (YCL and YCR), VII (YGL), XII (YLR), and XIV (YNL and YNR). In addition to the homologous ORFs, homologs to the two flanking *A. gossypii* tRNA genes (open squares), one snRNA gene (black squares), and the centromeres (black circles) were identified in the syntenic *S. cerevisiae* regions and connected by lines. Black rectangles represent two of 262 *A. gossypii* ORFs that lack a homolog in *S. cerevisiae* (table S3). The transcription direction is indicated by white arrows

and is identical in the syntenic *S. cerevisiae* ORFs. Thin horizontal lines above and below *S. cerevisiae* genes mark the extent of ancient synteny, as inferred from the reconstruction of gene deletion and genome rearrangement events (fig. S4). A major difficulty rests in the assignments of the ends of these lines in cases of extensive gene losses, e.g., between *YNL001W* and *YNL007C*. These cases are rare, and the final decision is based on the reconstruction of the most likely rearrangement events that interrupted the original gene order. End points of such events are marked by white arrows when the gene order in one of the two syntenic *S. cerevisiae* regions is affected (single breaks of synteny) and by black arrows when both gene orders are simultaneously affected (double breaks of synteny).

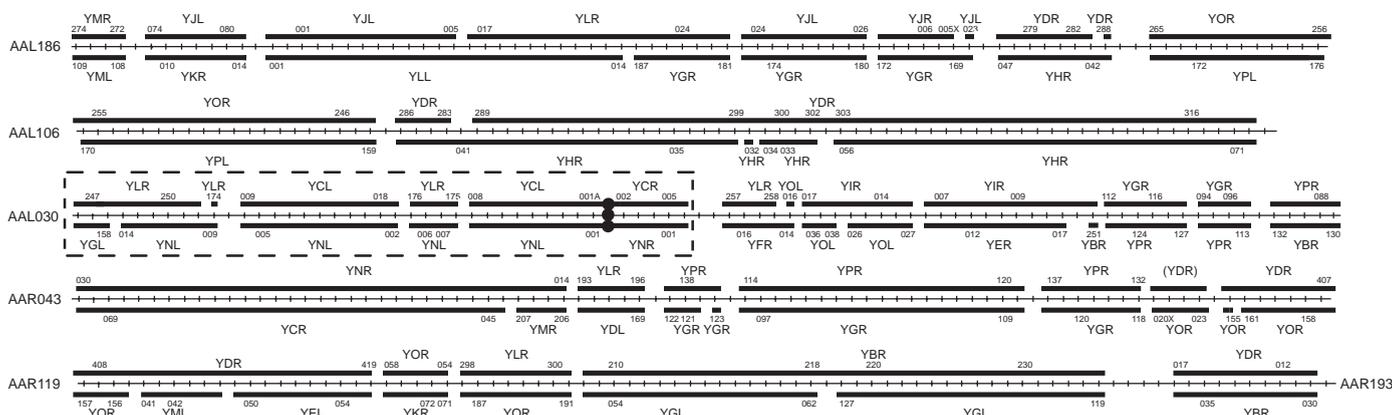


Fig. 3. Alignments of syntenic *S. cerevisiae* genes with chromosome I of *A. gossypii*. The vertical bars represent 381 ORFs, 27 tRNA genes, and 2 snRNA genes of *A. gossypii* chromosome I. Thick horizontal bars correspond to those regions of the *S. cerevisiae* genome that display synteny to the *A. gossypii* genes. The three capital letters and the three-digit numbers refer to the systematic nomenclature of *S. cerevisiae* genes. For example, the two bars marked as YJL 074 to 080 and YKR 010 to 014 (top left) encompass seven *A. gossypii* genes with homology, conserved gene order, and conserved gene orientation to the *S. cerevisiae* genes. This group of genes represents a cluster of ancient synteny, with ancient synteny referring to the degree of synteny at the time of *S. cerevisiae*

genome doubling, when most *A. gossypii* genes had two syntenic *S. cerevisiae* homologs. In order to match the ancient synteny, thick bars are often extended at one end to complement the loss of one of the duplicated *S. cerevisiae* genes. This ancient synteny map displays double breaks and single breaks of synteny. The different origins of these breaks in the boxed region are explained in detail in fig. S4. The gaps of the homology bars at double and single breaks are due to *A. gossypii* genes that either lack homology to any *S. cerevisiae* gene or display homology to a nonsyntenic *S. cerevisiae* gene. Centromeres are indicated by black circles and show the evolutionary conservation of the centromere region after genome doubling.

stream of *AAL174C* and *ACR029C* (Table 1 and table S3). These genome alignments are interrupted by 328 double breaks of synteny and 168 single breaks of synteny.

The essentially complete coverage of the seven *A. gossypii* chromosomes by clusters of ancient synteny, each containing two *S. cerevisiae* gene regions, demonstrates that both organisms originate from the same ancestor with seven or eight chromosomes. A speciation event, probably involving translocations (17) and an accompanying change in chromosome number, generated the precursors of *A. gossypii* and *S. cerevisiae*. At some later time, a genome duplication in the *S. cerevisiae* precursor opened new possibilities for functional divergence not available for the evolution of *A. gossypii*. The duplication event created ~5000 twin ORFs in the duplicated *S. cerevisiae* genome, and 496 of these ancient twin ORFs can still be seen in the double synteny patterns (table S7). Several of these twin ORFs diverged and now encode proteins of different functions like *ORC1*, which is essential for DNA replication, and *SIR3*, which is important for gene silencing (18). Other examples can be extracted from the functional descriptions in table S7. For 59 pairs of the twin ORFs, functions are not known.

What does the frequency of different types of synteny breaks tell us about the time span since *A. gossypii* and *S. cerevisiae* diverged? On the basis of adjusted numbers of synteny breaks (fig. S5), we estimate 120 viable genome rearrangements in the *A. gossypii* lineage and 180 viable rearrangements in the *S. cerevisiae* lineage (~60 before genome duplication). If one assumes similar rates of genome rearrangements in both species and takes into account a recent increase in *S. cerevisiae* rearrangement due to spreading of transposable elements (19), the time span since divergence of both species is about twice as long as the time span since the genome duplication in *S. cerevisiae*. This method for estimating relative evolutionary time scales from genome rearrangement frequencies has the potential to be used more often in the future, when additional whole genome synteny patterns become available.

Note added in proof: An *A. gossypii* genome browser is available at <http://agd.unibas.chl>. A recent publication by Kellis, Birren, and Lander (20) also demonstrates an ancient genome duplication of the *S. cerevisiae* genome.

References and Notes

1. K. P. Stahmann, J. L. Revuelta, H. Seulberger, *Appl. Microbiol. Biotechnol.* **53**, 509 (2000).
2. M. C. Wright, P. Philippsen, *Gene* **109**, 99 (1991).
3. S. Steiner, J. Wendland, M. C. Wright, P. Philippsen, *Genetics* **140**, 973 (1995).
4. Y. Ayad-Durieux, P. Knechtle, S. Goff, F. Dietrich, P. Philippsen, *J. Cell Sci.* **113**, 4563 (2000).

5. J. Wendland, Y. Ayad-Durieux, P. Knechtle, C. Rebeschung, P. Philippsen, *Gene* **242**, 381 (2000).
6. C. Alberti-Segui, F. Dietrich, R. Altmann-Johl, D. Hoepfner, P. Philippsen, *J. Cell Sci.* **114**, 975 (2001).
7. J. Wendland, P. Philippsen, *Genetics* **157**, 601 (2001).
8. P. Knechtle, F. Dietrich, P. Philippsen, *Mol. Biol. Cell* **14**, 4140 (2003).
9. R. Altmann-Johl, P. Philippsen, *Mol. Gen. Genet.* **250**, 69 (1996).
10. P. Philippsen *et al.*, *Nature* **387** (suppl.), 93 (1997).
11. K. H. Wolfe, D. C. Shields, *Nature* **387**, 708 (1997).
12. J. Wendland *et al.*, *Curr. Genet.* **35**, 618 (1999).
13. V. Wood *et al.*, *Nature* **415**, 871 (2002).
14. A. Goffeau *et al.*, *Science* **274**, 546 (1996).
15. J. E. Galagan *et al.*, *Nature* **422**, 859 (2003).
16. D. Lalo, S. Stettler, S. Mariotte, P. P. Slonimski, P. Thuriaux, *C. R. Acad. Sci. III* **316**, 367 (1993).
17. D. Delneri *et al.*, *Nature* **422**, 68 (2003).
18. S. P. Bell, J. Mitchell, J. Leber, R. Kobayashi, B. Stillman, *Cell* **83**, 563 (1995).
19. G. Fischer, S. A. James, I. N. Roberts, S. G. Oliver, E. J. Louis, *Nature* **405**, 451 (2000).
20. M. Kellis, B. W. Birren, E. S. Lander, *Nature*, in press; published online 7 March 2004 (10.1038/nature02424).

21. We thank Y. Bauer, A. Binder, K. Gaudenz, S. Goff, J. Hoheisel, M. Jacquot, P. Knechtle, C. Rebeschung, H.-P. Schmitz, J. Wendland, and the Syngenta (formerly Novartis) sequencing facility at Research Triangle Park for their assistance with this project, and A. Gladfelder, M. Primig, and R. Rischatsch for critical reading of the manuscript. Supported by the University of Basel (P.P.), Syngenta Biotechnology Inc. (formerly Novartis Agribusiness Biotechnology Research Inc.), Clemson University, and Duke University's Young Investigator Start-Up Fund (F.D.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1095781/DC1

Materials and Methods

Tables S1 to S8

Figs. S1 to S5

References and Notes

20 January 2004; accepted 24 February 2004

Published online 4 March 2004;

10.1126/science.1095781

Include this information when citing this paper.

Neuronal Activity Related to Reward Value and Motivation in Primate Frontal Cortex

Matthew R. Roesch* and Carl R. Olson

In several areas of the macaque brain, neurons fire during delayed-response tasks at a rate determined by the value of the reward expected at the end of the trial. The activity of these neurons might be related to the value of the expected reward or to the degree of motivation induced by expectation of the reward. We describe results indicating that the nature of reward-dependent activity varies across areas. Neuronal activity in orbitofrontal cortex represents the value of the expected reward, whereas neuronal activity in premotor cortex reflects the degree of motivation.

In numerous areas of the brain extending from the limbic system to the motor system, neuronal activity varies according to the size of the reward for which a monkey is working (1–15). Reward-dependent activity commonly has been viewed as representing the value of the goal for which the monkey is working; however, it might alternatively be related to the monkey's degree of motivation. Anticipation of a more valued reward leads to stronger motivation, as evidenced by measures of arousal, attention, and intensity of motor output (16–18).

On the assumption that motivated behavior depends on influences arising in the limbic system and acting on the motor system (19), we hypothesized that neuronal signals representing reward value predominate in the limbic system, whereas signals

reflecting the degree of motivation predominate in the motor system. To test this hypothesis, we recorded from two areas in which neurons exhibit robust reward-related activity: the orbitofrontal division of limbic cortex (OF) (Fig. 1) and the postarcuate premotor cortex (PM) (Fig. 1). OF plays an important role in motivated behavior (20–22). Its neurons respond to cues predicting the availability of foodstuffs at a rate determined by their appetitive or aversive value (13–14). PM is a region of high-order motor cortex (23–25). Its neurons fire during the delay period of an ocular delayed-response task at a rate determined by the direction of the impending saccade and by the size of the expected reward (9).

To achieve a dissociation between activities dependent on reward value and on motivation, we recorded from single neurons while two monkeys performed a task in which the degree of motivation was controlled independently by the magnitude of the reward promised in the event of success and the magnitude of the penalty threatened in the event of failure (26). On each trial, two cues

Center for the Neural Basis of Cognition, Mellon Institute, Room 115, 4400 Fifth Avenue, Pittsburgh, PA 15213, USA, and Department of Neuroscience, University of Pittsburgh, Pittsburgh, PA 15260, USA.

*To whom correspondence should be addressed. E-mail: roesch@cnbc.cmu.edu

ERRATUM

post date 23 April 2004

REPORTS: "The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome" by F. S. Dietrich *et al.* (9 April 2004, p. 304). An incorrect URL for the *Ashbya gossypii* genome browser was given in the "Note added in proof" on p. 307. The correct URL is <http://agd.unibas.ch/>.