

12. C. S. Sikes, A. P. Wheeler, A. Wierzbicki, A. S. Mount, R. A. Dillaman, *Biol. Bull.* **198**, 50 (2000).
13. L. H. Jodrey, *Biol. Bull.* **104**, 398 (1953).
14. K. M. Wilbur, L. H. Jodrey, *Biol. Bull.* **103**, 269 (1952).
15. K. M. Wilbur, in *Physiology of Mollusca*, K. M. Wilbur, C. M. Yonge, Eds. (Academic Press, New York, 1964), vol. 1, pp. 242–282.
16. A. F. Eble, in *The Eastern Oyster, Crassostrea virginica*, V. S. Kennedy, R. I. E. Newell, A. F. Eble, Eds. (Maryland Sea Grant College, College Park, MD, 1996), pp. 271–298.
17. T. C. Cheng, in *Invertebrate Blood Cells*, N. A. Ratcliffe, A. F. Rowley, Eds. (Academic Press, London, 1981), pp. 233–300.
18. E. Gosling, *Bivalve Molluscs, Biology, Ecology and Culture* (Blackwell Publishing, Oxford, 2003).
19. T. C. Cheng, in *The Eastern Oyster, Crassostrea virginica*, V. S. Kennedy, R. I. E. Newell, A. F. Eble, Eds. (Maryland Sea Grant College, College Park, MD, 1996), pp. 299–333.
20. M. Justus, S. Patel, A. S. Mount, poster presented at the 6th International Conference on Shellfish Restoration, Charleston, SC, 20 to 24 November 2002.
21. M. G. McCormick-Ray, T. Howard, *J. Invertebr. Pathol.* **58**, 219 (1991).
22. P. M. Hine, *Fish Shellfish Immunol.* **9**, 367 (1999).
23. Materials and methods are available as supporting material on Science Online.
24. M. R. Carrier, R. E. Palmer, R. S. Prezant, *Proc. Natl. Shellfish Assoc.* **70**, 139 (1980).
25. M. R. Carrier, in *The Eastern Oyster, Crassostrea virginica*, V. S. Kennedy, R. I. E. Newell, A. F. Eble, Eds. (Maryland Sea Grant College, College Park, MD, 1996), pp. 75–167.
26. P. S. Galtsoff, *Fish Bull.* **64**, 1 (1964).
27. K. M. Wilbur, A. S. M. Saleuddin, in *The Mollusca*, K. M. Wilbur, Ed. (Academic Press, New York, 1983), vol. 4, part 1, pp. 236–288.
28. S. E. Ford, S. A. Kanaley, D. T. Littlewood, *J. Invertebr. Pathol.* **61**, 49 (1993).
29. A. P. Wheeler, P. L. Blackwelder, K. M. Wilbur, *Biol. Bull.* **148**, 472 (1975).
30. N. Watabe, in *The Mollusca*, K. M. Wilbur, Ed. (Academic Press, New York, 1983), vol. 4, part 1, pp. 289–317.
31. R. J. Kingsley, *Am. Zool.* **24**, 833 (1984).
32. A. H. Scheltema, *Biol. Bull.* **184**, 57 (1993).
33. A. H. Scheltema, C. Schander, *Biol. Bull.* **198**, 121 (2000).
34. We thank E. Ruppert, R. Montanucci, and D. Bushek for careful reading and suggestions that have improved the quality of this manuscript. We also thank J. Hatchell for his efforts in cell enumeration and construction of the oyster holding facility; B. Kay for technical help with EM; J. Mount for assistance in revised figure preparation; and T. McNutt Scott, M. Johnstone, S. Patel, and M. Justus for invaluable discussions. We gratefully acknowledge support from NSF under award no. 20020445.

Supporting Online Material
www.sciencemag.org/cgi/content/full/304/5668/297/DC1
 Materials and Methods
 Figs. S1 to S5
 References

15 August 2003; accepted 2 March 2004

Characterization of a Common Susceptibility Locus for Asthma-Related Traits

Tarja Laitinen,^{1,2} Anne Polvi,² Pia Rydman,¹ Johanna Vendelin,^{1,2} Ville Pulkkinen,¹ Paula Salmikangas,¹ Siru Mäkelä,² Marko Rehn,¹ Asta Pirskanen,¹ Anna Rautanen,⁴ Marco Zucchelli,⁵ Harriet Gullstén,^{2,5} Marina Leino,⁶ Harri Alenius,⁶ Tuula Petäys,⁷ Tari Haahtela,⁷ Annika Laitinen,³ Catherine Laprise,⁹ Thomas J. Hudson,¹⁰ Lauri A. Laitinen,⁸ Juha Kere^{*2,5}

Susceptibility to asthma depends on variation at an unknown number of genetic loci. To identify susceptibility genes on chromosome 7p, we adopted a hierarchical genotyping design, leading to the identification of a 133-kilobase risk-conferring segment containing two genes. One of these coded for an orphan G protein-coupled receptor named GPRA (G protein-coupled receptor for asthma susceptibility), which showed distinct distribution of protein isoforms between bronchial biopsies from healthy and asthmatic individuals. In three cohorts from Finland and Canada, single nucleotide polymorphism-tagged haplotypes associated with high serum immunoglobulin E or asthma. The murine ortholog of GPRA was up-regulated in a mouse model of ovalbumin-induced inflammation. Together, these data implicate GPRA in the pathogenesis of atopy and asthma.

Asthma is a complex phenotype with a proven genetic component, and several projects to map susceptibility genes for asthma and re-

lated traits have been undertaken (1). The first published genome-wide scan in asthma suggested six tentative genetic loci, among them chromosome 7p, which was then strongly implicated in a study of Finnish and Canadian families and confirmed in West Australian families (2–4). To positionally clone asthma candidate genes on chromosome 7p, we studied the Finnish Kainuu subpopulation and considered three alternative hypotheses. First, that only one copy of the susceptibility allele may have survived in this population, with long-conserved haplotypes being observed, as would be consistent with some previous findings (5–7). Second, there might exist a founder effect, common to many European populations. This would be consistent with the common disease/common variant hypothesis, as observed for example

in psoriasis (8, 9). In this case, the carrier frequency should be higher than a few percent, with only a short conserved haplotype (<200 kb) detected (8). Third, numerous mutations might exist in the putative gene, in which case only weak or absent haplotype associations might be detectable.

To distinguish between these hypotheses, we adopted a genotyping scheme whereby we increased the density of markers used, with intermediate analyses to guide further genotyping (Fig. 1A). More specifically, if genetic association analysis suggested that a haplotype occurred in patients more often than in controls, additional markers were genotyped to either exclude or support the identity-by-descent of the haplotypes observed in unrelated patients. For the haplotypes to be identical by descent, all newly typed markers would have to be identically shared between them. The genotyping was done on 86 original genome scan families and an additional 103 trios (all together, 874 subjects) (10). Successive rounds of genotyping and analysis by the haplotype pattern mining (HPM) algorithm (11) suggested the strong association of a conserved haplotype pattern spanning between NM51 and SNP563704, separated by 46 kb (Fig. 1A). The HPM algorithm searches for allele patterns shared between several haplotypes among large sets of unrelated haplotypes.

To fully explore the genetic variation in associated haplotypes, we sequenced nonrepetitive DNA segments in this interval (from position 506,401 to 638,799 in the public sequence NT_000380; all positions are given with reference to this sequence) in one patient homozygous for the susceptibility haplotype and one control subject homozygous for the most common (nonrisk) haplotype. These sequences were then compared to the public sequence (NT_000380). Two observations emerged from these analyses: first, the patient did not reveal a single instance of heterozygosity, confirming the identity-by-descent of this chromosome segment. Second, comparison of the susceptibility

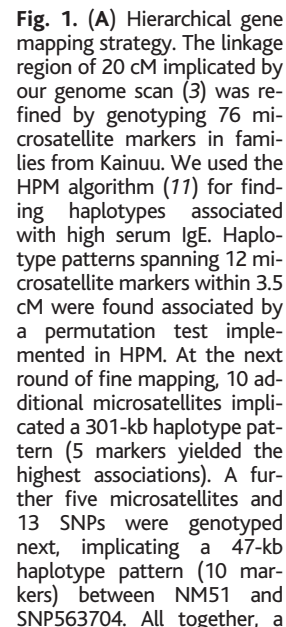
¹GeneOS Limited, 00251 Helsinki, Finland. ²Department of Medical Genetics, ³Department of Anatomy, Biomedicum Helsinki, University of Helsinki, 00014 Helsinki, Finland. ⁴Finnish Genome Center, University of Helsinki, 00014 Helsinki, Finland. ⁵Department of Biosciences at Novum and Clinical Research Centre, Karolinska Institutet, 14157 Huddinge, Sweden. ⁶Department of Industrial Hygiene and Toxicology, Finnish Institute of Occupational Health, 00251 Helsinki, Finland. ⁷Department of Allergological Diseases, ⁸Department of Medicine, Helsinki University Central Hospital, 00029 Helsinki, Finland. ⁹Community Genomic Medicine Centre, Université du Québec à Chicoutimi, Saguenay, Quebec G7H 2B1 Canada. ¹⁰McGill University and Genome Quebec Innovation Centre, Montreal, Quebec H3A 1A4 Canada.

*To whom correspondence should be addressed. E-mail: juha.kere@biosci.ki.se

To determine the limits of the critical region, we genotyped additional SNPs in 131 trios [304 high immunoglobulin E (IgE) associated and 220 control chromosomes], yielding a total of 51 SNPs between 490,331 and 691,245 base pairs (bp) and an average marker density of 4 kb (tables S2 and S3). Analysis of the data by HPM revealed strong association of a conserved 133-kb pattern between 514,743 and 647,327 bp (Fig. 1A). Within this segment, there was strong linkage disequilibrium between the markers (Fig. 1B). A permutation test for association showed $P \leq 0.01$ for all 43 markers in this 133-kb segment (10,000 permutations). For

We subsequently genotyped two additional population samples with subjects that had been diagnosed either for asthma (from Northeastern Quebec, Canada) or high IgE (from North Karelia, Finland) and corresponding family-based controls. We used 22 SNPs among Quebec families (514 asthma-associated and 258 control chromosomes) and 29 SNPs among North Karelian families (75 high IgE-associated and 49 control chromosomes). A haplotype pattern with the same limits as in Kainuu was also identified in Quebec (Fig. 1C) and North Karelia (12). Most of the SNPs were shared, although some were distinct from those in Kainuu. In the three populations, 13 SNPs across the most conserved 77 kb formed seven alternative haplotypes with frequencies $>2\%$. We next sequenced the 133-kb segment from six

We next tested the hypothesis that the related haplotypes (identified either on the basis of high IgE in Kainuu or North Karelia or asthma in Quebec) together conferred risk in all three populations. The risk haplotypes could be tagged by SNP522363 (allele C, table S1) and indeed associated with significant risk ($P = 0.004$ for all data combined, all three populations contributing), consistent with the common disease/common variant hypothesis (8). The relative risk for high se-



133-kb region was sequenced around this segment from a homozygous patient with asthma. Eighty polymorphisms were identified by comparison to the public genomic sequence. **(B)** Plot displaying pairwise linkage disequilibrium (LD) values between 51 markers genotyped in Kainuu trios. The markers are arranged in linear order along both axes starting from lower left corner. The scale from blue to dark red indicates LD ranging from 0 to 1, respectively. LD values shown for patients and controls are essentially identical. A segment of very strong LD between 41 markers is detected. In the lower panel, the same markers were evaluated for association to high serum IgE by a permutation test. **(C)** Plot displaying LD in French Canadian trios between 22 markers spanning the same segment as in **(B)**. A segment of strong LD has the same boundaries as in Kainuu. In the lower panel, permutation test indicated significant association to asthma in French Canadian trios. **(D)** Phylogenetic analysis of haplotypes H1 to H7

within a 77-kb segment in Kainuu, North Karelia, and Quebec. The same seven haplotypes occur in all three populations at frequencies >2%. H4 and H5 are the most common risk-associated haplotypes in Kainuu, H7 in North Karelia, and H2 among French Canadians. H1, H3, and H6 are nonrisk haplotypes in all three populations.

Fig. 2. Gene content around the conserved 133-kb haplotype segment (gray box). **(A)** The 133-kb segment spans from intron 2 to intron 5 of GPRA. GPRA undergoes alternative splicing with multiple variants; the three longest variants are shown (thin lines joining exons marked E1 to E9b). Exon 2 donor site may join to alternative exon 3 acceptor sites, separated by 33 bp in the same reading frame, and there are two alternative 3' exons, 9a and 9b. Further splice variants may skip exon 3 or 4 or both, suggesting an involvement of the associated polymorphisms in regulation of splicing and protein isoform production. **(B)** In the opposite DNA strand, there is a previously unknown gene, AAA1, with at least 18 exons (numbered 1 to 18) with complex alternative splicing. AAA1 spans a total of 500 kb of genomic sequence. Eight exons of GPRA (E1 to E8) are shown for orientation. **(C)** Northern blot hybridization with a 1285-bp full-length GPRA-A cDNA probe (left) and a mixed splice variant probe for AAA1 (right). A 2.4-kb transcript is visible in all nine lanes (upper arrow) and a 1.8-kb transcript (lower arrow) in four tissues for GPRA. Several alternative transcripts are seen for AAA1 (arrows).

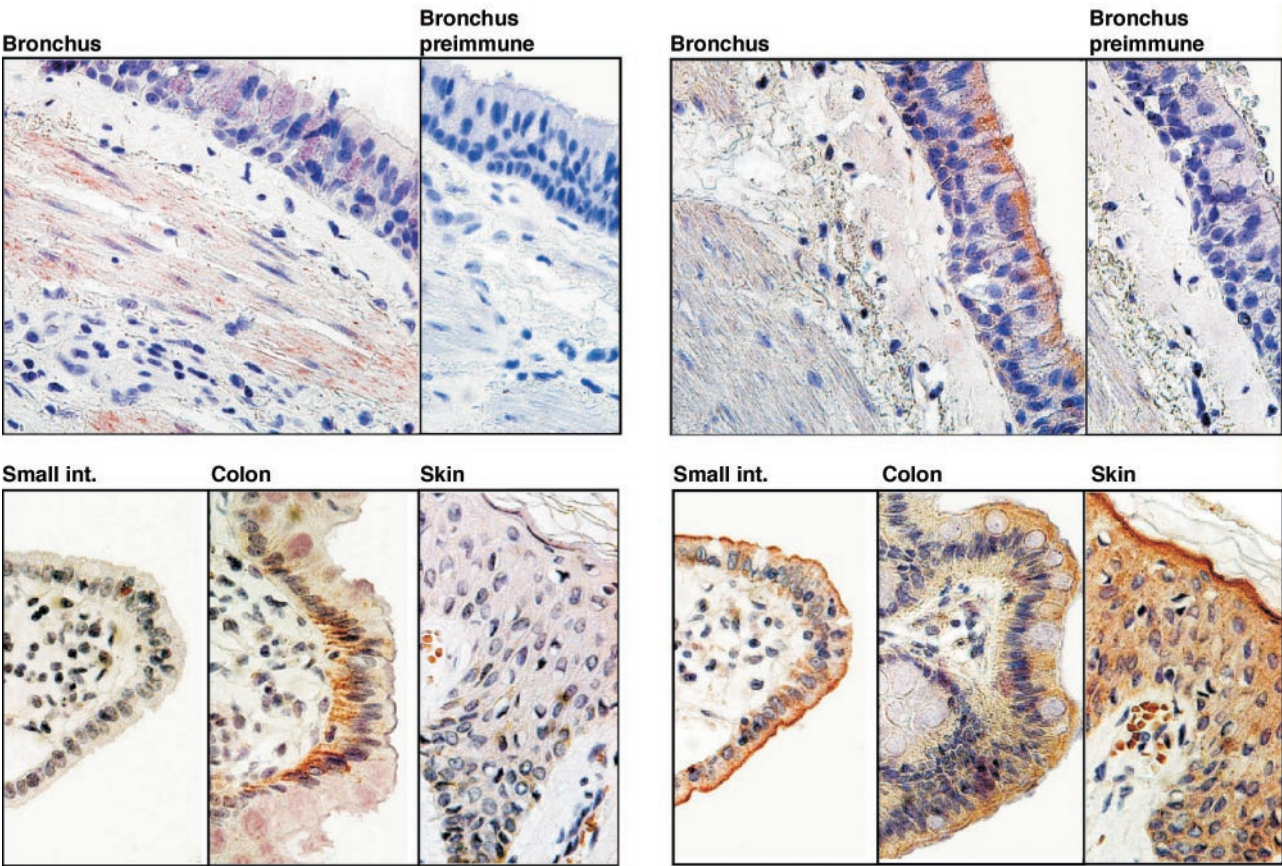
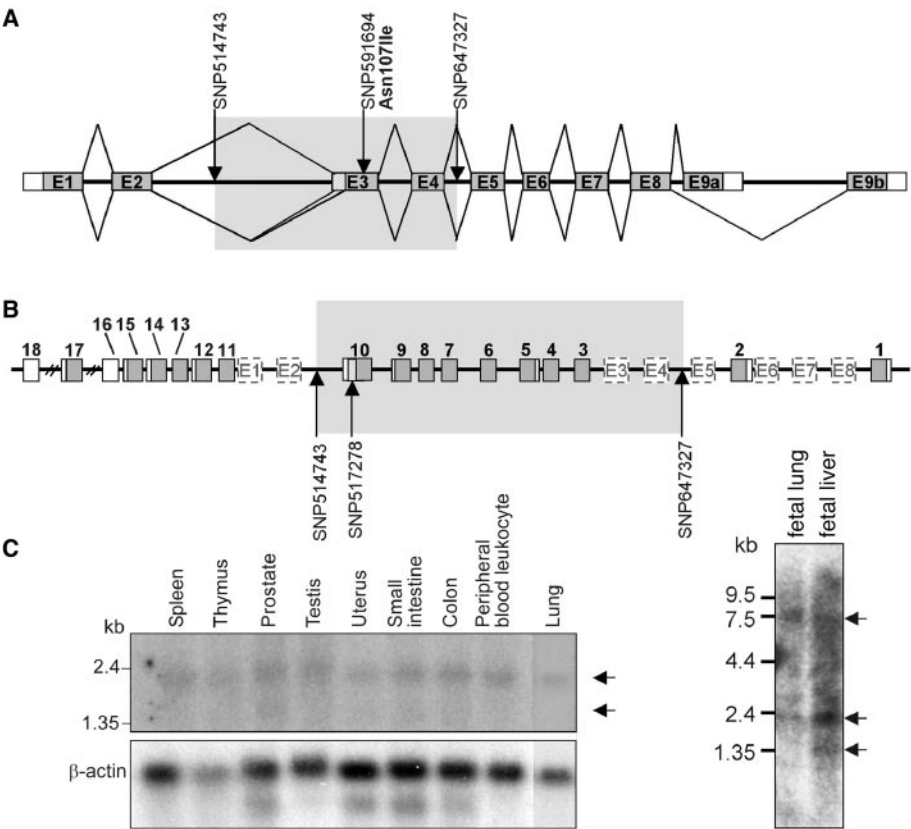


Fig. 3. Expression of the A (left) and B (right) isoforms of GPRA as assessed by immunohistochemistry. The A isoform is found in bronchial smooth muscle cells, basally in colon epithelium, and in occasional basal

keratinocytes. The B isoform stains apical epithelial cells in bronchus and gut and all layers of the epidermis. Negative controls are shown with preimmune sera.

rum IgE among H4 or H5 carriers in Kainuu was 1.4 (95% confidence interval 1.1 to 1.9, $P = 0.01$), and for asthma among homozygous H2 carriers in Quebec, 2.5 (95% confidence interval 2.0 to 3.1, $P = 0.0009$). Corresponding transmission disequilibrium test yielded $P = 0.05$ for Kainuu families ($n = 86$ trios). To assess whether genetic linkage to chromosome 7p could be explained by these haplotypes, we considered parent-offspring transmissions and sibling-pair sharing of high IgE in Kainuu families (3). One of the risk haplotypes cosegregated in 26 of 51 transmissions (51%) and was shared in 26 of 40 sibling pairs (65%), suggesting that a majority of the linkage signal was because of the observed risk haplotypes.

These results strongly implicated the 133-kb genomic segment as a susceptibility locus for asthma-related phenotypes. To understand more precisely how the observed genetic vari-

ation might influence susceptibility, we examined the DNA segment for specific genes (10). Two genes were identified, one with exons 3 to 5 and the other with exons 3 to 10 lying within the susceptibility haplotype. Structures of both indicated complex alternative splicing of the mRNAs (Fig. 2 and tables S4 and S5), suggesting translation to varying protein isoforms. One of the genes (Fig. 2A) was predicted to belong to the G protein-coupled receptor family, and we named this gene GPRA (for G protein-coupled receptor for asthma susceptibility). The two main transcripts of GPRA (A and B) had alternative 3' exons encoding proteins of 371 and 377 amino acids, respectively (GenBank AY310326 and AY310327). The sequences of all predicted isoforms of the other gene, named AAA1 (for asthma-associated alternatively spliced gene 1, GenBank AY312365 to AY312373) (Fig. 2B) showed only weak homologies to any previously identified proteins.

Transcripts for both genes were detected by Northern blot hybridizations (Fig. 2C). Both genes displayed coding polymorphisms in the asthma susceptibility haplotype. In GPRA, SNP591694 changed an amino acid (Asn¹⁰⁷ → Ile¹⁰⁷) in the first exoloop lining the putative ligand-binding pocket. Even if the position was functionally important, it was unclear how the altered protein would affect cells and tissues.

Polyclonal antibodies were raised against the different carboxyl termini of the A and B isoforms of the predicted GPRA protein (10). Immunohistochemical staining of bronchus, gut, and skin sections showed that the A isoform is predominantly expressed by smooth muscle cells, whereas the B isoform was predominantly detected in epithelial cells (Fig. 3). In bronchial biopsies, the isoform patterns were distinct between asthma patients and control samples (Fig. 4A). Most clearly, strong expression of the B isoform in smooth muscle cells in asthmatic airways compared with an absence of such staining in control samples. Staining for the B isoform in epithelial cells varied between healthy individuals (Figs. 3 and 4A) but was consistently stronger in the asthma samples than in controls. The A isoform showed no consistent differences. These results suggested also that one or more SNPs or DIPs in the risk haplotypes might critically alter the balance between the isoforms.

We examined next the potential role of the mouse ortholog of GPRA (*Gpra*) in a mouse model of ovalbumin-induced lung inflammation (10). In general agreement with the results in human asthma, *Gpra* mRNA was significantly up-regulated in mouse lung after ovalbumin tests in sensitized compared with nonsensitized mice (Fig. 4B). No ortholog for AAA1 was found in mouse and could thus not be assayed. These data support a role for GPRA in the pathogenesis of asthma and provide an animal model that may be useful to further assess the functions of this protein.

In contrast to the GPRA results, several lines of evidence suggested that AAA1 may not represent a protein-coding gene, although its expression was modified by the haplotype (Fig. 4C). Its longest open reading frame comprised only 74 potential amino acids, and in vitro translation failed to yield a stable polypeptide. Transiently transfected cells did not produce recombinant protein. Polyclonal peptide antibodies detected the antigen but no proteins in Western blots or immunohistochemistry (12).

Recently, a small number of candidate genes influencing susceptibility to asthma have been identified by positional cloning. However, the biochemical mechanisms linking the candidate genes to pathogenetic processes in asthma remain poorly understood (1, 13–15). The properties of GPRA make it a strong candidate for involvement in the pathogenesis of asthma and other IgE-mediated diseases, as well as a possible drug target. GPRA might act as a receptor for an unidentified ligand. The putative ligand, iso-

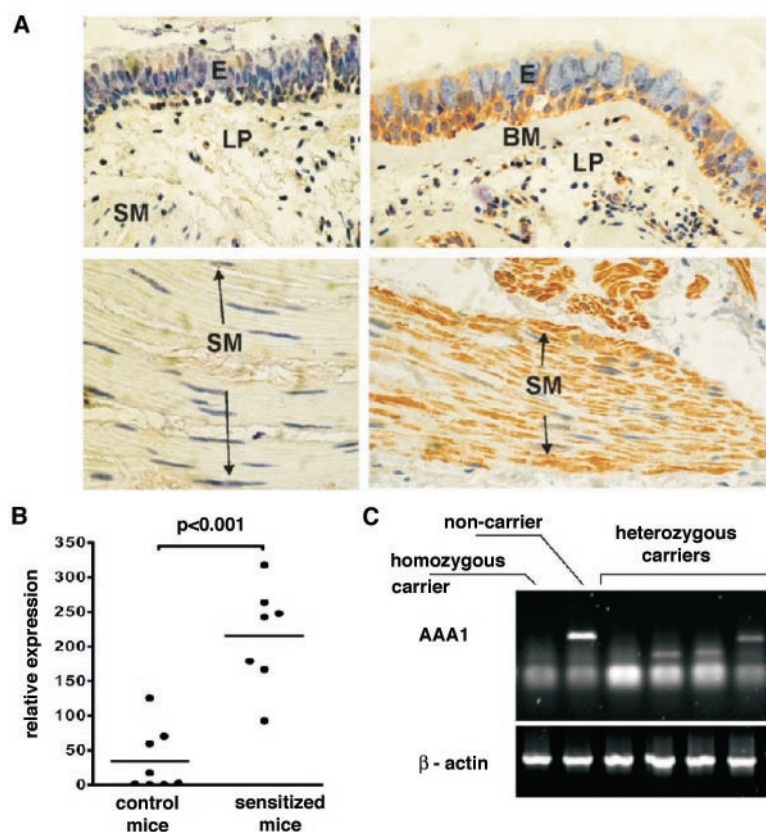


Fig. 4. (A) Expression of GPRA isoform B in bronchial biopsies from a healthy control (left) and an asthma patient (right). E, epithelium; BM, basement membrane; LP, lamina propria; SM, smooth muscle. (Top) The airway epithelium in the control sample shows only faint staining. In the asthma patient, the epithelium shows Goblet cell hyperplasia and basement membrane thickening typical of asthma and positive immunostaining for GPRA throughout the ciliated cells but not Goblet cells. (Bottom) The asthmatic smooth muscle stains strongly positive for GPRA isoform B, in contrast to the negative finding in control. Results are typical of 8 asthmatic and 10 control biopsies studied. (B) Relative expression levels of *Gpra* mRNA in lungs from sensitized ($n = 7$) and control ($n = 8$) mice after inhaled ovalbumin challenge. *Gpra* was significantly up-regulated in sensitized compared with control mice. (C) Variable alternative splicing for AAA1 depending on genotype. Reverse transcription polymerase chain reaction spanning exons 6 to 10b of AAA1 was performed on lymphoblast RNA samples genotyped for the susceptibility haplotype. Only noncarriers process normal amount of the exon 6–10b transcript, whereas homozygote and heterozygotes show either absent transcript or smaller splice variants. Beta-actin was used as control in parallel amplifications.

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