

Sequencing the *Aspergillus fumigatus* genome

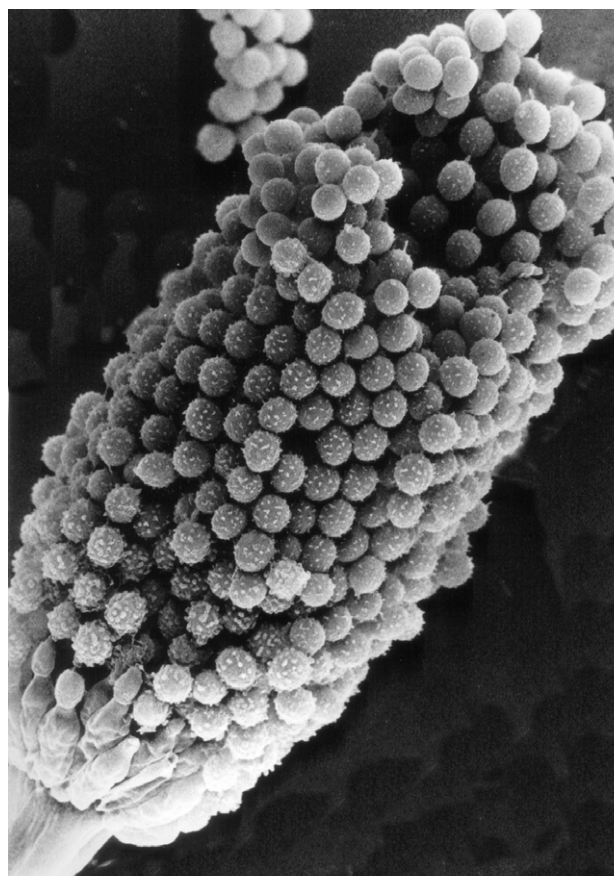
David W Denning, Michael J Anderson, Geoffrey Turner, Jean-Paul Latgé, and Joan W Bennett

Aspergillus fumigatus is the most common mould pathogen of human beings and unusually causes both invasive disease in immunocompromised patients and allergic disease in patients with atopic immune systems. 4% of patients dying in modern European teaching hospitals have invasive aspergillosis and it is the leading infectious cause of death in leukaemia and bone marrow transplant patients. Until 2001, only two licensed antifungal drugs were available to treat aspergillosis—amphotericin B and itraconazole. Its 28–30 Mb genome is being sequenced in an international collaboration, with the Wellcome Trust Sanger Institute (UK) and The Institute for Genomic Research (TIGR, USA) as the two main centres. A whole-genome shotgun approach was adopted and initiated in 2001 with an expected completion date in 2003. The complete sequence will permit identification of pathways specific to pathogenic *Aspergillus* species, help identify new targets for antifungal drugs, and enable investigations into the basic biology of fungi. Numerous secondary metabolic pathways with biotechnological applications and pharmacological properties are found in the Aspergilli and the genome sequence will facilitate research in this area.

Lancet Infect Dis 2002; **2**: 251–53

Aspergillus is a remarkable genus of microorganisms. Among the 182 recognised species,¹ *Aspergillus fumigatus* (figure) is the most common human and animal pathogen^{2,3} and can be the most common thermophilic species found in the air in human dwellings.⁴ The sequencing of the genome was started in 2001 and is likely to be completed by the beginning of 2003, in an international effort involving the Wellcome Trust Sanger Institute (UK), The Institute for Genomic Research (TIGR, USA), the University of Salamanca (Spain), Complutense University (Spain), the Centro de Investigaciones Biológicas, Madrid, Spain, and the Pasteur Institute (France). Involvement has been determined primarily by new funding opportunities (eg, National Institutes of Health) and decisions, and facilitated by international collaborative meetings. Costs of sequencing are falling as the project proceeds, but will not exceed \$6.5 million including finishing and curatorship. The size of the genome is estimated to be 28–30 Mb.

Aspergillus was first recognised as an organism in 1729 by Micheli.⁵ The first described cases of infection were in a jackdaw (*Corvus monedula*) in 1815⁶ and in a person from Edinburgh in 1842.⁷ *A. fumigatus* as a species was delineated by Fresenius in 1863.⁸ Allergic disease of the lungs was



Conidial head of *A. fumigatus* as seen by scanning electron microscopy. Each conidium has a diameter of 3–5 μm .

recognised by Popoff in 1887⁹ and was mainly an occupational disease of wig combers. The first invasive fatal case (in 1890) was in a 3-year-old child with aspergillus airways disease, without evident immune compromise.¹⁰ It was not until 1953 that the pathologist Rankin, working in Gloucester, UK, described invasive aspergillosis in a patient

DWD and MJA are at the School of Medicine, University of Manchester, Manchester, UK; GT is at the Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, UK; J-PL is at the Laboratoire d'*Aspergillus*, Department of Structure and Dynamics of Genomes, Pasteur Institute, Paris, France; and JWB is at the Department of Cell and Molecular Biology, Tulane University, New Orleans, LA, USA.

Correspondence: Dr David Denning, Room 1.800, Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK. Tel +44 (0)161 291 5818; fax +44 (0)161 291 5806; email ddenning@man.ac.uk

who died with aplastic anaemia as a consequence of chloramphenicol toxicity.¹¹ Today, less than 50 years later, 4% of patients dying in modern European teaching hospitals have invasive aspergillosis,^{12,13} and it is the leading infectious cause of death in leukaemia and bone-marrow-transplant patients. In the USA in 1996, the cost of treating diagnosed cases of invasive aspergillosis was estimated at \$633 million, with each case costing approximately \$64 500.¹⁴ Until 2001, only two antifungal drugs were available to treat aspergillosis—amphotericin B, licensed in the early 1960s, and itraconazole, licensed in the early 1990s.

A fumigatus is a remarkable and unusual pathogen. In a recent classification of all human pathogens, *Aspergillus* species were singled out in their own group as being capable of causing infection at both extremes of immune function.¹⁵ Highly immunocompromised patients are at great risk of life-threatening invasive disease, which diminishes in frequency and severity as immune function improves.² By contrast, those with fully functional immune systems do not get invasive infection, but if atopic, may get allergic disease due to aspergillus.¹⁶ These allergic entities are known as allergic bronchopulmonary aspergillosis and eosinophilic fungal sinusitis. Local lung damage, as after tuberculosis, also predisposes to aspergillus disease of the lung (aspergilloma and chronic necrotising pulmonary aspergillosis), often leading eventually to respiratory failure.¹⁶ The rising frequency of tuberculosis worldwide will increase in tandem the number of such aspergillosis cases. One of the reasons *A fumigatus* is capable of causing these diseases in so many people, is that it is an ubiquitous organism in the air and distributed world-wide. Why it is by far the most common mould pathogen and fungal allergen is one of the questions the genome sequencing project will assist in answering. It is clear that *A fumigatus*, above all other mould fungi, has evolved specific features responsible for its pathobiology.¹⁷

The natural niche of the thermotolerant *A fumigatus* is decaying organic material; as a consequence, it is one of the most common inhabitants of compost.⁴ Many of the other species of *Aspergillus* are also of concern to human society. On the positive side, *Aspergillus nidulans* has been one of a handful of model organisms that have contributed hugely to our understanding of classical genetics.^{18,19} It differs from *A fumigatus* in having a sexual stage, which is amenable to meiotic analysis. It is a rare human pathogen.²⁰ *Aspergillus niger* is one of a small number of industrial fungi and is used for the production of citric acid, enzymes, and the heterologous expression of various proteins²¹, as well as being an occasional pathogen of man.²² *Aspergillus oryzae* (a domesticated clone of *Aspergillus flavus*) is used extensively in northeast Asia for the production of sake, miso, bean curd, and other fermented foods.²¹ On the negative side, *A flavus* is remarkable in being both a human and plant pathogen, and, to a lesser degree than its close relative *Aspergillus parasiticus*, is a producer of aflatoxin, a major mycotoxin in foodstuffs.²³ Certain other species are thought to be significant allergens and producers of mycotoxins,^{23,24} including *Aspergillus clavatus*²⁵ and *Aspergillus versicolor*²⁶. *Aspergillus terreus* is distinguished by its relative frequency in certain hospitals in the USA and

Austria and because it is intrinsically amphotericin B resistant.²⁷ It is the major cause of disseminated aspergillosis in German shepherd dogs.²⁸ Little is known about the ecological niche, pathogenicity to other life forms, and secondary metabolite production of most of the other *Aspergillus* species.

Initially the sequencing strategy for the genome of *A fumigatus* was to be a clone-by-clone-based approach. A simpler and less expensive approach would be division of labour by chromosome between the multiple sequencing centres, but this approach was not possible because the chromosomes of *A fumigatus* could not be fully separated. A whole genome shotgun approach has been adopted, because of improvements in assembly capability to accommodate mid-sized genomes, good working relationships between the groups involved, and a strong management structure. All assemblies over 1 kb and end sequences from the BAC library clones are being made publicly available on the TIGR and Sanger websites, before submission to GenBank and EMBL. Detailed discussions concerning annotation software and its “training” are in progress between the two major sequencing partners for several mid-sized eukaryotic genomes (particularly *Plasmodium falciparum*), which has been fostered by the international collaborations necessary to garner enough resource and expertise to complete each genome.

Comparison with non-virulent *Aspergillus* strains or species will identify pathways specifically found in the pathogenic species, which will help determine why *A fumigatus* is the most common human mould pathogen and capable of such extremes of infection from allergic to invasive disease. Direct identification of repeated sequences, such as inactive transposons and microsatellites, that are useful for strain typing, will allow a better understanding of the epidemiology of aspergillosis. Understanding microaerophilic growth adaptation will be possible, as well as a comprehensive analysis of resistance of *A fumigatus* to both the immune response and antifungal drugs. New drug targets will emerge for the medical and agricultural markets, which are urgently needed given the poor responses to current drugs and the existence of drug-resistant isolates.²⁷ Early diagnosis is pivotal for the improvement of outcome in serious aspergillus disease and new leads will be provided by the complete genome sequence.

In addition to its intrinsic interest, numerous benefits will accrue from the genome sequence of *A fumigatus* and associated large-scale experiments. Along with data from the two other moulds that have been sequenced in their entirety and the data made available in the public domain (*Neurospora crassa* and *Phanerochaete chrysosporium*), the genetic machinery controlling the two characteristic morphologies of all moulds and no other eucaryotes—namely hyphal extension and conidial production—will be available for scrutiny. Evolutionary relationships will become clearer and will establish the frequency of mammalian-like proteins in lower eukaryotes, such as the insulin-like proteins.²⁹ Numerous secondary metabolic pathways with biotechnological applications and pharmacological properties are found in the Aspergilli.

One therapeutic example includes the lipopeptides or echinocandins derived in 1974 from a fermentation broth of *A. nidulans* var *echinulatus*.³⁰ The most advanced of the echinocandins (Merck's caspofungin) has just been granted a licence by the US Food and Drug Administration for, of all indications, the treatment of refractory aspergillosis. Additional therapeutic examples include the angiogenesis-inhibitor fumagillin from *A. fumigatus* (analogues of which have provided a novel treatment for microsporidiosis)³¹ and lovastatin, (a successful cholesterol-lowering agent)³² produced by *A. terreus*. Other secondary metabolites produced by *A. fumigatus* include gliotoxin, mannitol, phthioic acid, fumitremorgin, verruculogen, fumigaclavine, helvolic acid, and

shingofungin.²⁴ *A. fumigatus* presumably has some of the enzymatic apparatus to produce aflatoxin,³³ since it produces an intermediate in the pathway perhaps in a similar manner to *A. nidulans*³⁴ rather than *A. oryzae*, which only contains unexpressed genes of this pathway.³⁵ The synthetic pathways of other less well understood compounds with pharmacological or other properties, such as conidial diffusates that inhibit macrophage phagocytosis,³⁶ will be tractable for the first time.

It is our hope that the understanding of the biological reasons that turn a saprophytic fungus into a major opportunistic pathogen will come from the information derived from the first complete genome sequence of a human pathogenic mould.

References

- Pitt JI, Samson RA, Frisvad JC. List of accepted species and their synonyms in the family Trichocomaceae. In: Samson RA, Pitt JI, eds. Integration of modern taxonomic methods for *Penicillium* and *Aspergillus* classification. Amsterdam: Harwood Publishers, 2000: 9–49.
- Denning DW. Invasive aspergillosis. *Clin Infect Dis* 1998; 26: 781–805.
- Smith JMB. Opportunistic mycoses of man and other animals. Wallingford, UK: CAB International, 1989.
- Latgé JP. *Aspergillus fumigatus* and aspergillosis. *Clin Microbiol Rev* 1999; 12: 310–50.
- Micheli PA. Nova plantarum genera juxta Tournefortii methodum disposita. Florence: Bernardo Paperini, 1729.
- Mayer AC. *Deutsch Arch Physiol* 1815 (Meckel). Cited by Rénon.
- Bennett JH. On the parasitic vegetable structures found growing in living animals. *Trans R Soc Edinburgh* 1842; 15: 277–79.
- Fresenius G. Beiträge zur Mykologie. Frankfurt: HL Bronner, 1850, 81.
- Popoff LW. Ein fall von mycosis aspergillina nebst einigen Bemerkungen über aenliche Erkrankungen der Respirationsweg. Varsovie 1887; 3: 316.
- Wheaton SW. Case primarily of tubercle, in which a fungus (*aspergillus*) grew in the bronchi and lung, stimulating actinomycosis. *Path Trans* 1890; 41: 34–37.
- Rankin NE. Disseminated aspergillosis and moniliasis associated with agranulocytosis and antibiotic therapy. *BMJ* 1953; 1: 918–19.
- Groll AH, Shah PM, Mentzel C, Schneider M, Just-Neubling G, Huebner K. Trends in the postmortem epidemiology of invasive fungal infections at a university hospital. *J Infect* 1996; 33: 23–32.
- Vogeser M, Haas A, Aust D, Ruckdeschel G. Postmortem analysis of invasive aspergillosis in a tertiary care hospital. *Eur J Clin Microbiol Infect Dis* 1997; 16: 1–6.
- Dasbach EJ, Davies GM, Teutsch SM. Burden of aspergillosis-related hospitalisations in the United States. *Clin Infect Dis* 2000; 31: 1524–28.
- Casadevall A, Pirofski L-A. Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. *Infect Immun* 1999; 67: 3703–313.
- Denning DW. Chronic forms of pulmonary aspergillosis. *Clin Microbiol Infect* 2001; 7 (suppl 2): 25–31.
- Latgé J-P. The pathobiology of *Aspergillus fumigatus*. *Trends Microbiol* 2001; 9: 382–89.
- Pontecorvo G, Roper JA, Hemmons LM, MacDonald KD, Bufton AWJ. *Adv Genet* 1953; 5: 141–238.
- Martinelli SD, Kinghorn JR, eds., *Aspergillus: 50 years on*. Amsterdam: Elsevier, 1994.
- Segal BH, DeCarlo ES, Kwon-Chung KJ, Malech HL, Gallin JI, Holland SM. *Aspergillus nidulans* infection in chronic granulomatous disease. *Medicine* 1998; 77: 345–54.
- Bennett JW, Klich MA. *Aspergillus*. In: Flickinger MC, Drew SW, eds., *Encyclopaedia of bioprocess technology: fermentation, biocatalysis, and bioprocessing*. New York: Wiley & Sons, 1999.
- Severo LC, Geyer GR, da Silva Porto N, Wagner MB, Londero AT. Pulmonary *Aspergillus niger* colonisation. Report of 23 cases and a review of the literature. *Rev Iberoam Micol* 1997; 14: 104–10.
- Pitt JI. Toxigenic fungi: which are important? *Med Mycol* 2000; 38 (suppl 1): 17–22.
- Moss MO. Mycotoxins of *Aspergillus* and other filamentous fungi. *Soc Appl Bacteriol Symp Ser* 1989; 18: 69S–87S.
- Blyth W, Grant IW, Blackadder ES, Greenberg M. Fungal antigens as a source of sensitisation and respiratory disease in Scottish maltworkers. *Clin Allergy* 1977; 7: 549–62.
- Jarvis JQ, Morey PR. Allergic respiratory disease and fungal remediation in a building in a subtropical climate. *Appl Occup Environ Hyg* 2001; 16: 380–88.
- Moore CB, Sayers N, Mosquero J, Slaven J, Denning DW. Antifungal drug resistance in *Aspergillus*. *J Infect* 2000; 41: 203–20.
- Kelly SE, Shaw SE, Clark WT. Long-term survival of four dogs with disseminated *Aspergillus terreus* infection treated with itraconazole. *Aust Vet J* 1995; 72: 311–13.
- Le Roith D, Shiloach J, Roth J, Lesniak MA. Evolutionary origins of vertebrate hormones: substances similar to mammalian insulins are native to unicellular eukaryotes. *Proc Natl Acad Sci USA* 1980; 77: 6184–88.
- Nyfelers R, Keller-Schierlein W. Metabolites of microorganisms. 143. Echinocandin B, a novel polypeptide-antibiotic from *Aspergillus nidulans* var. *echinulatus* isolation and structural components. *Helv Chim Acta* 1974; 57: 2459–77.
- Coyle C, Kent M, Tanowitz HB, Wittner M, Weiss LM. TNP-470 is an effective antimicrosporidial agent. *J Infect Dis* 1998; 177: 515–18.
- Kennedy J, Auclair K, Kendrew SG, Park C, Vederas JC, Hutchinson CR. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* 1999; 284: 1368–72.
- Trail F, Mahanti N, Rarick M, et al. Physical and transcriptional map of an aflatoxin gene cluster in *Aspergillus parasiticus* and functional disruption of a gene involved early in the aflatoxin pathway. *Appl Environ Microbiol* 1995; 61: 2665–73.
- Brown DW, Yu JH, Kelkar HS, et al. Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. *Proc Natl Acad Sci USA* 1996; 93: 1418–22.
- Watson AJ, Fuller LJ, Jeenes DJ, Archer DB. Homologs of aflatoxin biosynthesis genes and sequence of *aflR* in *Aspergillus oryzae* and *Aspergillus sojae*. *Appl Environ Microbiol* 1999; 65: 307–310.
- Robertson MD, Seaton A, Raeburn JA, Milne LJ. Inhibition of phagocyte migration and spreading by spore diffusates of *Aspergillus fumigatus*. *J Med Vet Mycol* 1987; 25: 389–96.

Useful websites

<http://www.aspergillus.man.ac.uk>
http://www.sanger.ac.uk/Projects/A_fumigatus/
<http://www.tigr.org/tdb/e2k1/afu1/>