

# Draft genome sequences of *Neurospora crassa* clade B, isolated from burned *Cytisus* sp. plants in France

Lucas Bonometti,<sup>1</sup> Fabien De Bellis,<sup>2,3</sup> Sandrine Cros-Arteil,<sup>1</sup> Elise Gueret,<sup>4</sup> Pierre Gladieux<sup>1</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 2.

**ABSTRACT** *Neurospora crassa* clade A is a model system for genetics, biochemistry, molecular biology, and experimental evolution. Here, we present the draft genome sequences of four isolates of *N. crassa* clade B. These data represent a valuable resource to investigate the population biology and evolutionary history of *N. crassa sensu lato*.

**KEYWORDS** *Neurospora*, ascomycetes, endophytes

*Neurospora crassa* is “the model microbe that opened the book of genetics at the chapter on biochemistry” (1). Despite its wide use as a model organism, central aspects of the biology and evolutionary history of *N. crassa* remain poorly understood (2). For instance, the life cycle, fine-scale genetic structure, and history of divergence among the three lineages of *N. crassa* [referred to as clades A, B, and C (3)] are not fully understood. *N. crassa* clade A isolate OR74A was the first filamentous fungus to be fully sequenced (4). Since the release of the reference genome, whole genome data for approximately 60 additional isolates have been published, mostly representing clade A (5–7). The public release of high-quality genomic data representative of the diversity of *N. crassa* is needed to fill the gaps in our knowledge of its natural history.

Four strains of *N. crassa* were collected as wood chips with sporulating bright orange colonies from burnt *Cytisus* sp. plants after a wildfire in Villeveyrac, France, in 2017. Samples Vill-A1-2 and Vill-A1-3 originated from the same shrub, while samples Vill-B4 and Vill-C1-2 originated from two other shrubs. Each wood chip was placed in a separate paper envelope and stored at –20°C (8). Single-spore isolation was carried out by depositing 100 µL of sterile water onto a mycelial colony and spreading the suspension in a Petri dish with water agar (2%) and chloramphenicol (200 g/mL). After incubating 12 hours at room temperature, a germling was removed using a needle, transferred to a Potato Dextrose Agar medium covered with a cellophane jam jar cover (Hutchinson, Roubaix, France), and incubated 36 hours on the bench before harvesting mycelium.

Genomic DNA was extracted from mycelium harvested from cellophane membranes using a protocol strictly identical to that described in Thierry et al. (9). DNA was fragmented by sonication (Covaris, Massachusetts). DNA libraries were prepared with TruSeq nano kits and sequenced using NovaSeq 6000 (Illumina, California) with 150 nucleotide paired-end reads and 500 base-pair insert size. The number of read pairs ranged between 8,006,852 and 9,635,233. Read quality was controlled using TRIMMOMATIC v0.32 (10).

Genomes were assembled using ABYSS 2.2.1 (11) with default parameters. The *k*-mer size was varied from 60 to 130, so as to maximize the N50. Optimal *k*-mer sizes and genome assembly statistics are reported in Table 1. Genome completeness estimated with BUSCO 5.4.3 [9] (sordariomycetes\_odb10 set) was above 97%. Bacterial contamination was estimated using KRAKEN2 v2.1.2 (default settings) with the PLUSFPF database (12).

Isolates were assigned to *N. crassa* clades based on their evolutionary distance from reference genomes OR74A [clade A; PRJNA132, PRJNA13841 (4)], FGSC4830 (clade B;

**Editor** Jason E. Stajich, University of California, Riverside, Riverside, California, USA

Address correspondence to Pierre Gladieux, pierre.gladieux@inrae.fr.

The authors declare no conflict of interest.

See the funding table on p. 3.

**Received** 14 July 2023

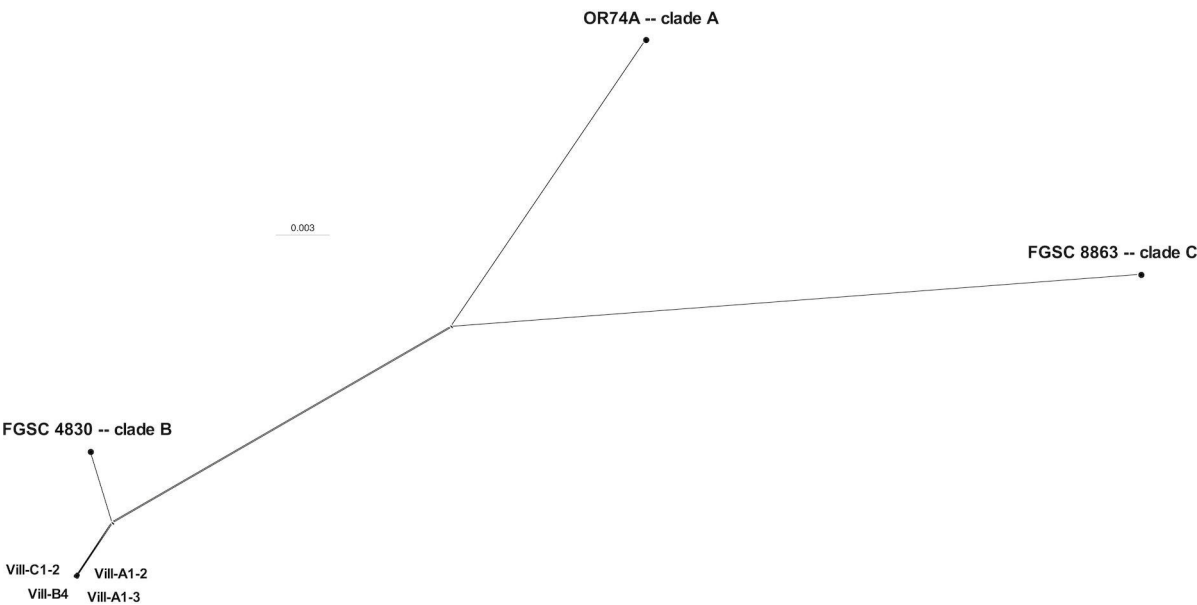
**Accepted** 4 October 2023

**Published** 9 November 2023

Copyright © 2023 Bonometti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

**TABLE 1** Assembly quality statistics and *k*-mer sizes used for assembly of four isolates of *N. crassa* clade B isolated from burned *Cytisus* sp. plants in France

Parameters	Isolates			
	Vill-A1-2	Vill-A1-3	Vill-B4	Vill-C1-2
<i>k</i> -mer size	92	84	84	78
Number of scaffolds	6,746	9,156	7,567	9,218
L50	22	19	21	20
N50	512,892	528,255	481,520	551,850
Assembly length	37,345,753	37,773,353	37,603,129	36,795,690
Busco score (%)	97.5	97.1	97.2	97.4
GC content (%)	50.73	50.63	50.56	51.19
Bacterial contamination (% reads)	0.14	0.08	0.14	0.04



**FIG 1** Neighbor-net phylogenetic network estimated with SPLITSTREE v4 showing the assignment of four isolates of *N. crassa* collected from burnt *Cytisus* sp. plants in Villeveyrac (France) to *N. crassa* clade B. The network was inferred using Hamming distance calculated from 182,553 polymorphic sites identified in the concatenated sequences of 3,540 single-copy core genes (6,397,566 base pairs in total). A phylogenetic network was constructed, instead of a phylogenetic tree, because evolution is not expected to be tree-like at the intraspecific level for a fungus whose populations are recombining.

PRJNA371206), and FGSC8863 [clade C (7, 13)]. Coding sequences at 3,540 single-copy core genes resulting from the BUSCO analysis (14) were aligned with TRANSLATORX 1.1 (15) to maintain the coding frame [aligner: PRANK v170427 (16); parameter -b5 = half in GBLOCKS 0.91b (17)]. Alignments were concatenated (6,397,566 bp), and isolates were assigned to *N. crassa* clade B using a NeighborNet network built with SPLITSTREE v4.018.2 (18) (default settings; Fig. 1).

ACKNOWLEDGMENTS

MGX acknowledges financial support from France Génomique national infrastructure, funded as part of “Investissement d’Avenir” program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09).

AUTHOR AFFILIATIONS

<sup>1</sup>PHIM Plant Health Institute, University of Montpellier, INRAE, CIRAD, Institut Agro, Montpellier, France  
<sup>2</sup>CIRAD, UMR AGAP Institut, Montpellier, France

<sup>3</sup>AGAP Institut, University of Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

<sup>4</sup>MGX-Montpellier GenomiX, University of Montpellier, CNRS, INSERM, Montpellier, France

## AUTHOR ORCIDs

Pierre Gladieux  <http://orcid.org/0000-0003-1929-1576>

## FUNDING

Funder	Grant(s)	Author(s)
France Génomique	ANR-10-INBS-09	Elise Gueret

## AUTHOR CONTRIBUTIONS

Lucas Bonometti, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review and editing | Fabien De Bellis, Conceptualization, Investigation | Sandrine Cros-Arteil, Investigation, Validation | Elise Gueret, Investigation, Validation, Writing – review and editing | Pierre Gladieux, Conceptualization, Project administration, Supervision, Writing – original draft, Writing – review and editing

## DATA AVAILABILITY

The draft genome sequences and the raw read data were deposited in DDBJ/ENA/GenBank under BioProject accession number [PRJEB64279](https://www.ncbi.nlm.nih.gov/bioproject/PRJEB64279).

## REFERENCES

- Davis RH, Perkins DD. 2002. Timeline: *Neurospora*: a model of model microbes. *Nat Rev Genet* 3:397–403. <https://doi.org/10.1038/nrg797>
- Gladieux P, De Bellis F, Hann-Soden C, Svedberg J, Johannesson H, Taylor JW. 2020. *Neurospora* from natural populations: population genomics insights into the life history of a model microbial eukaryote. *Methods Mol Biol* 2090:313–336. [https://doi.org/10.1007/978-1-0716-0199-0\\_13](https://doi.org/10.1007/978-1-0716-0199-0_13)
- Dettman JR, Jacobson DJ, Taylor JW. 2003. A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. *Evolution* 57:2703–2720. <https://doi.org/10.1111/j.0014-3820.2003.tb01514.x>
- Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma L-J, Smirnov S, Purcell S, et al. 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422:859–868. <https://doi.org/10.1038/nature01554>
- Villalba de la Peña M, Summanen PAM, Liukkonen M, Kronholm I. 2022. Chromatin structure influences rate and spectrum of spontaneous mutations in *Neurospora crassa*. *bioRxiv*. <https://doi.org/10.1101/2022.03.13.484164>
- Svedberg J, Vogan AA, Rhoades NA, Sarmarajeewa D, Jacobson DJ, Lascoux M, Hammond TM, Johannesson H. 2021. An introgressed gene causes meiotic drive in *Neurospora sitophila*. *Proc Natl Acad Sci U S A* 118:e2026605118. <https://doi.org/10.1073/pnas.2026605118>
- Svedberg J, Hosseini S, Chen J, Vogan AA, Mozgova I, Hennig L, Manitchotpisit P, Abusharekh A, Hammond TM, Lascoux M, Johannesson H. 2018. Convergent evolution of complex genomic rearrangements in two fungal meiotic drive elements. *Nat Commun* 9:4242. <https://doi.org/10.1038/s41467-018-06562-x>
- Perkins DD, Turner BC, Barry EG. 1976. Strains of *Neurospora* collected from nature. *Evolution* 30:281–313. <https://doi.org/10.1111/j.1558-5646.1976.tb00910.x>
- Thierry M, Charriat F, Milazzo J, Adreit H, Ravel S, Cros-Arteil S, Borron S, Sella V, Kroj T, loos R, Fournier E, Tharreau D, Gladieux P. 2022. Correction: maintenance of divergent lineages of the rice blast fungus *Pyricularia oryzae* through niche separation, loss of sex and post-mating genetic incompatibilities. *PLoS Pathog* 18:e1010944. <https://doi.org/10.1371/journal.ppat.1010944>
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, Jahesh G, Khan H, Coombe L, Warren RL, Birol I. 2017. ABySS 2.0: resource-efficient assembly of large genomes using a bloom filter. *Genome Res* 27:768–777. <https://doi.org/10.1101/gr.214346.116>
- Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. *Genome Biol* 20:257. <https://doi.org/10.1186/s13059-019-1891-0>
- Svedberg J, Hosseini S, Chen J, Vogan AA, Mozgova I, Hennig L, Manitchotpisit P, Abusharekh A, Hammond TM, Lascoux M, Johannesson H. 2018. Convergent evolution of complex genomic rearrangements in two fungal meiotic drive elements. *Nat Commun* 9:4242. <https://doi.org/10.1038/s41467-018-06562-x>
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res* 38:W7–13. <https://doi.org/10.1093/nar/gkq291>
- Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. *Methods Mol Biol* 1079:155–170. [https://doi.org/10.1007/978-1-62703-646-7\\_10](https://doi.org/10.1007/978-1-62703-646-7_10)
- Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23:254–267. <https://doi.org/10.1093/molbev/msj030>