BMC Research Notes

RESEARCH NOTE

Open Access

Improved genome assembly of *Onygena corvina*, a keratin-degrading fungus

Siddhi Pavale[1], Vincent G. H. Eijsink[1] and Sabina Leanti La Rosa[1*]

Abstract

Objective *Onygena corvina* is a non-pathogenic, saprophytic fungus that colonizes feathers, hooves, and hair, and represents a valuable source of keratin-degrading enzymes. The only genome assembly of *O. corvina* available to date was obtained for the strain CBS 281.48 using Illumina short-read sequencing, yielding a reference genome composed of 521 contigs with a contig N50 of 0.229 Mb.

Results Here, we report an improved *O. corvina* CBS 281.48 genome assembly generated using a high-quality hybrid approach that combines Illumina short-read and Oxford Nanopore long-read sequencing. The new assembly consists of only 13 contigs totaling 21.8 Mb, with an N50 of 4.4 Mb, and has a completeness of 98.4%. A total of 7,232 protein-coding genes were annotated using an integrative approach that combines *de novo* predictions, homology-based inferences, and RNA-sequencing–guided evidence. Notably, 158 putative protease-coding genes were identified representing a substantial increase from the 73 predicted proteases in the previous annotation. Our improved genome assembly and associated gene annotations will facilitate comparative genomics, and high-resolution mapping of transcriptomic and proteomic data, to advance research on fungal physiology and fungal abilities to degrade recalcitrant substrates such as keratin.

Keywords Onygena corvina, Genome assembly, Keratin degradation

Introduction

Onygena corvina is a saprotrophic ascomycete in the order Onygenales, capable of colonizing keratinous materials such as feathers, hooves, and hair [1, 2]. Its keratinolytic potential, coupled with a non-pathogenic nature, makes this fungus a promising candidate for keratin waste valorization [3, 4]. However, the only existing reference genome (*O. corvina* CBS 281.48) is fragmented and only partially annotated, limiting in-depth functional analyses [5, 6]. Here, we present a high-quality genome assembly for *O. corvina* CBS 281.48 and associated gene annotations, which were obtained using a hybrid approach combining Illumina short-read and Oxford Nanopore long-read sequencing for DNA as well as Illumina sequencing for RNA.

Genome assembly and annotation

O. corvina CBS 281.48 was obtained from the CBS-KNAW Collection of the Westerdijk Fungal Biodiversity Institute (Utrecht, Netherlands). For DNA and RNA extraction, mycelial biomass was generated in 500 ml flasks with 200 ml potato dextrose broth that were inoculated with five to eight mycelial plugs and incubated at 25 °C, 100 rpm for 7 days. Mycelia were harvested using miracoth filters (Merck KGaA, Darmstadt, Germany),

*Correspondence:
Sabina Leanti La Rosa
sabina.leantilarosa@nmbu.no
[1]Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway

washed with distilled water, flash-frozen in liquid nitrogen, and stored at − 80 °C until further use.

Genomic DNA was extracted using the ZYMO Research Quick-DNA Miniprep Plus Kit (Zymo Research, Freiburg, Germany). For long-read sequencing, a library was prepared with a Native Barcoding Expansion Kit (EXP-NBD104, Oxford Nanopore Technologies, Oxford, UK) and sequenced on a Nanopore PromethION 48 with a Flow Cell R10.4.1 (FLO-PRO114M). Library construction and sequencing were performed at Biomarker Technologies GmbH (Münster, Germany). Base-calling and adapter trimming were performed using Dorado v0.7.2 (Oxford Nanopore Technologies) in super-accurate mode. The long-read sequencing generated 380,647 reads, totaling approximately 2.33 Gb, with an average length of 6,114 bp and with an N50 of 7,952 bp. For short-read sequencing, a sample library was prepared for Illumina pair-end sequencing (2 × 150 bp) with the Nextera DNA Flex kit (Illumina, San Diego, USA) followed by sequencing with an Illumina NovaSeq X platform. Overall, the sequencing depth was 106 X.

Total RNA was extracted using the ZYMO Research Quick-RNA Miniprep Plus Kit following the manufacturer's instructions. The RNA-seq library was constructed using the Ultima Dual-mode mRNA Library Prep Kit for Illumina (Yeasen Biotechnology, Shanghai, China) following the manufacturer's instructions. The library was then sequenced on the Illumina NovaSeq X platform, yielding 6.77 Gb of 150-bp paired-end reads.

Table 1 Summary of predicted repeat sequences in the *O. corvina* genome, including the number of elements, their total length (bp), and their proportion (%) of the genome

Type	Number	Length (bp)	Percentage (%)
ClassI	161	203,372	0.93
ClassI/LINE	55	160,052	0.73
ClassI/LTR	3	465	0.00
ClassI/LTR/Copia	11	779	0.00
ClassI/LTR/Gypsy	90	42,214	0.19
ClassI/PLE|LARD	2	175	0.00
ClassII	55	14,548	0.07
ClassII/Crypton	4	480	0.00
ClassII/Helitron	2	122	0.00
ClassII/MITE	22	10,126	0.05
ClassII/TIR	21	3,376	0.02
ClassII/Unknown	6	499	0.00
SSR	147	240,272	1.10
Unknown	351	307,127	1.41
Total	363	702,375	3.22

Repeats are classified hierarchically as Class/Superfamily/Family. The first term indicates the transposon class (Class I = retrotransposon, Class II = DNA transposon), the second term indicates the superfamily, and the third term indicates the specific family. Abbreviations: Long Interspersed Nuclear Elements; LTR, Long Terminal Repeats; PLE, Penelope-like elements; MITE, Miniature Inverted-repeat Transposable Elements; TIR, Terminal Inverted Repeats

For hybrid *de novo* assembly, based-called ONT reads were first filtered by quality using FiltLong v0.2.1 (https://github.com/rrwick/Filtlong) with the min_length 1000 and keep_percent 90 parameters. Illumina reads were trimmed (trailing Q at least 28) and filtered (length < 50) with trimmomatic 0.38 [7]. The genome assembly was generated using Canu v2.3 [8] with the -nanopore-raw parameters and genome size was set as 30 Mb. The obtained assembly was polished using Nanopore reads with Racon v. 1.5.0 [9] (2 iterations). Next, the polished sequences were subjected to a final round of error correction using short-read Illumina data with Pilon v1.24 [10].

There were no assembly gaps, and quality metrics indicated 99.96% coverage. The coverage depth of the assembled genome was 66.1 X. Assessment of assembly completeness using BUSCO v6.0 [11] and the onygenales_odb12 dataset (downloaded in September 2025) identified 3674 of 3733 (98.4%) conserved onygenales genes, indicating a near-complete assembly. Repetitive sequences were identified using a species-specific repeat database constructed with LTR_FINDER [12], MITE-Hunter [13], RepeatScout [14], and PILER-DF [15]. This database was classified with PASTEClassifier [16], merged with Repbase [17], and used for annotation with RepeatMasker v4.1.5 [18]. The analysis revealed that 3.22% of the genome consisted of repeats and that simple sequence repeats and unclassified elements made up the bulk of repetitive DNA (Table 1). Retroelements (Class I) represented 0.93% of the genome, including Long Interspersed Nuclear Elements (0.73%, LINE) and Long Terminal Repeat elements (0.19%, LTR/Gypsy retroelements). Transposons (Class II) were detected at 0.07%, including Miniature Inverted-repeat Transposable Elements (MITE, 0.05%) and Terminal Inverted Repeat elements (TIR, 0.02%). In addition, 1.32% of the genome was represented by Simple Sequence Repeats (SSR). Some repetitive elements that could not be assigned to either Class I or Class II were also detected (1.41%).

Prediction of protein-coding genes was carried out by integrating *de novo*, homology-based, and transcriptome-supported approaches. Genscan [19], GlimmerHMM v3.0.4 [20], GeneID v1.4.5 [21], and SNAP [22] were used for *de novo* prediction without RNA-seq support. Augustus v3.1 [23] incorporated RNA-seq reads to improve exon-intron boundary predictions while GeMoMa v1.9 [24] was used to predict genes by leveraging homology with closely related fungal genomes (*Coccidioides immitis*, *Coccidioides posadasii* and *Ophidiomyces ophidiicola*) as well as RNA-seq evidence to refine the gene annotations. Transcriptomic data were mapped to the genome with HISAT2 v2.2.1 [25] and assembled with StringTie v2.2.3 [25], after which coding regions were identified using TransDecoder v5.7.1 [26] and PASA v2.5.3 [27],

using default parameters. The EVidenceModeler (EVM) v2.1.0 [28] was used to combine the predictions (de novo, homology, and transcriptome-based) of the individual software into a consensus gene list and refined further using PASA v2.5.3 [27] with RNA-seq transcripts. The process resulted in 7,232 protein-coding genes. The average gene length was 2164 bp, with 3.37 exons per gene and an average coding sequence length of 479 bp. Of note, 99.01% of gene predictions were supported by transcriptome or homology evidence, indicating high reliability of these predictions.

Comparative genomic analysis

Compared to the previous genome assembly (GCA_000812245.1), the updated *O. corvina* genome is more contiguous. While the total assembly length remains ~ 21.8 Mb, the number of contigs was reduced from 521 to 13. The contig N50 was increased substantially from 0.229 Mb to 4.41 Mb, with the largest contig now spanning 6.85 Mb (Table 2). Sequence alignment using Minimap2 [29] and DGENIES [30] revealed 98.65% identity with the earlier version of the genome, indicating high identity and collinearity (Fig. 1A).

In addition to the 13 contigs representing the nuclear genome, a separate contig (contig 6) corresponding to the mitochondrial genome was also recovered. While the length of this contig (21.4 Mb) is slightly less than the previous mitochondrial genome assembly (NC_082835.1, 24.3 Mb), sequence alignment with Minimap2 [29] and DGENIES [30] revealed high identity (99.92%) between the region spanning 9–8,599 bp of the new mitochondrial assembly and 15,753–24,343 bp of the old assembly, as well as between the region spanning 8,606–21,413 bp of the new assembly and 3–12,810 bp of the old assembly (Fig. 1B).

Table 2 Assembly statistics for the previous (PRJNA270018) and new (PRJNA1280007) version of the *O. corvina* genome

Characteristic	Previous version (PRJNA270018)	New version (PRJNA1280007)
Genome Size (Mb)	21.7	21.8
Number of contigs	521	13
Largest Contig size (Mb)	0.933	6.854
Contig N50 (Mb)	0.229	4.410
GC percent	48	48

Functional genomics and keratin degradation potential

We additionally conducted a secondary metabolite cluster analysis using antiSMASH (v8.0) for fungi, with the detection strictness set to relaxed [31]. A total of 32 biosynthetic gene clusters (BGCs) were identified, including terpene, type I polyketide synthase (T1PKS), non-ribosomal peptide synthetase (NRPS), and hybrid PKS-NRPS clusters. Several BGCs showed high similarity to known clusters: YWA1 [32] (polyketide); AbT1 [33] and chrysogine [34] (both NRPS); clavaric acid [35] (terpene); and tolypyridone C [36] (hybrid PKS-NRPS).

Given the ability of *O. corvina* to colonize and degrade keratin-rich substrates, we explored its keratinolytic potential through genome-wide prediction of proteases. Through motif recognition using the Homology to Peptide Pattern (Hotpep) pipeline, 158 proteases were predicted, a substantial expansion from the 73 identified

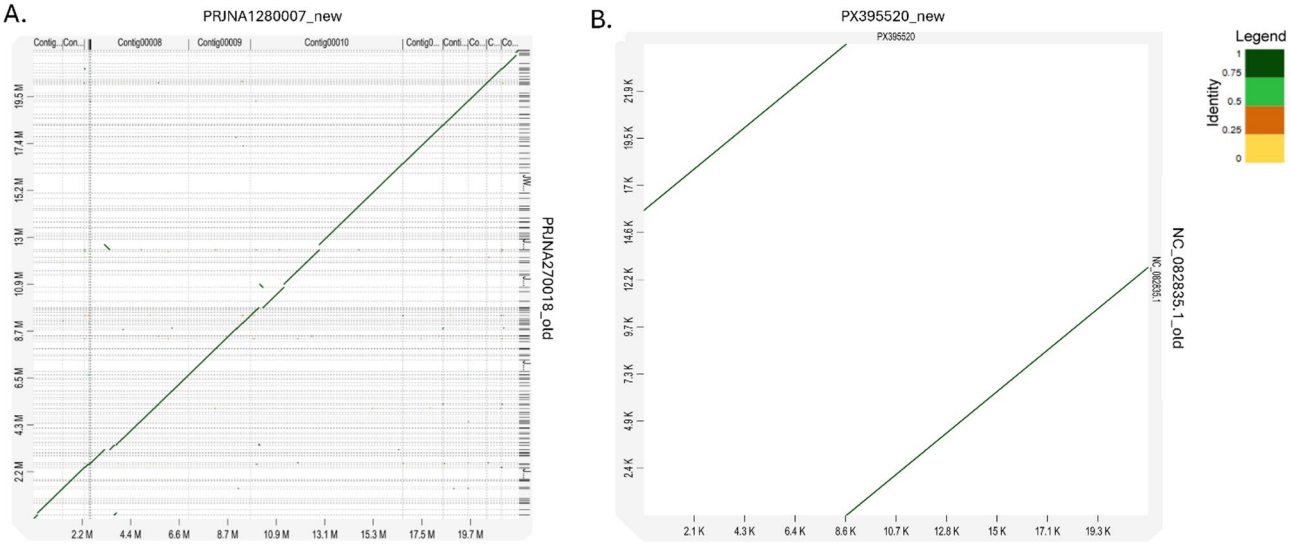

Fig. 1 Comparison of the new and previous version of the *O. corvina* nuclear genome and mitochondrion. **A** Alignment between the new version (PRJNA1280007) (top) and the previous version (PRJNA270018/GCA_000812245.1) (right) of the *O. corvina* nuclear genome. **B** Alignment between the new (PX395520) (top) and previous version (NC_082835.1) (right) of the *O. corvina* mitochondrial genome. The alignments were visualized with D-GENIES, where the coloration corresponds to the percent identity as per the legend

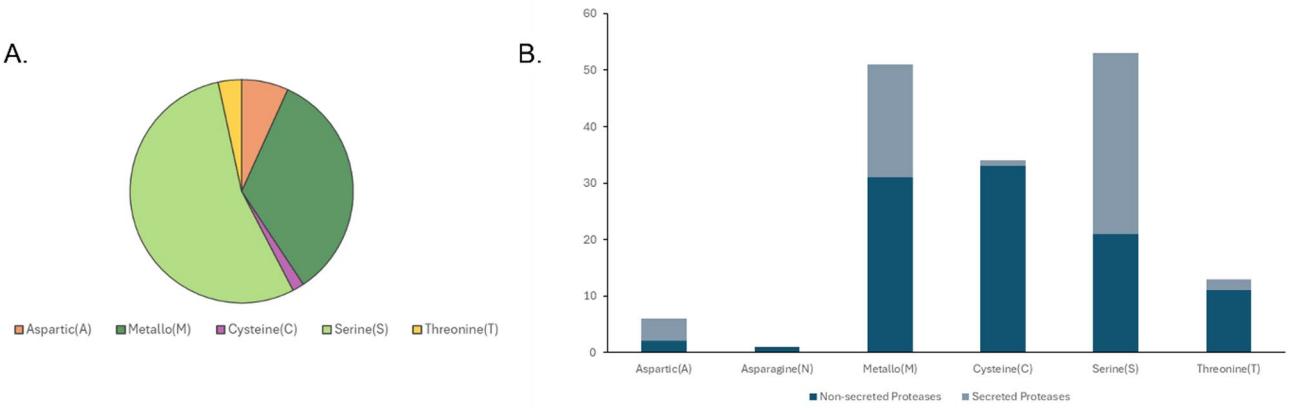

Fig. 2 MEROPS-based classification of proteases encoded in the *O. corvina* genome, categorized into serine (S), metalloprotease (M), threonine (T), cysteine (C), asparagine (N) and aspartic (A) families. **A** Distribution of the 59 putatively secreted proteases across these families. **B** Bar chart showing the comparative abundance of secreted versus non-secreted (intracellular and membrane-bound) proteases (158 enzymes in total). Light blue bars represent secreted proteases, while dark blue bars indicate non-secreted proteases

in the previous assembly [1]. These spanned across six MEROPS [37] families including aspartic proteases (A; n = 6), asparagine/peptidyl-lyases (N; n = 1), metalloproteases (M; n = 51), cysteine proteases (C; n = 34), serine proteases (S; n = 53), and threonine proteases (T; n = 13). Out of 158 proteases, 59 were predicted to be secreted based on signal peptide and transmembrane domain analysis using SignalP 5.0 [38] and TMHMM [39] v2.0 (Fig. 2A). Further classification based on subcellular localization showed that the secreted proteases were predominantly serine and metalloproteases, while the non-secreted proteases largely belonged to cysteine and metalloprotease families (Fig. 2B). Additionally, we identified 189 carbohydrate-active enzymes (CAZymes) in the *O. corvina* genome, using dbCAN3 [40]. Of these, 38 were predicted to be secreted, including glycoside hydrolases (GH; n = 27), carbohydrate esterases (CE; n = 2), auxiliary activity enzymes (AA; n = 8), and glycosyltransferases (GT; n = 1). Of note, two genes (*Onygena_corvina0G042260.1* and *Onygena_corvina0G066340.1*) coding for lytic polysaccharide monooxygenases (LPMOs) in AA11 family were detected, corroborating earlier findings for *O. corvina* and related keratin-degrading fungi [41]. While AA11-type LPMOs are known to act on crystalline chitin [42, 43] and soluble chitin fragments [44], these redox enzymes have been proposed to facilitate cleavage of keratin, possibly by targeting its carbohydrate components [2].

Limitations

Comparative genomic analyses beyond basic assembly comparisons were not performed, and further investigations are needed to determine the phylogenetic placement and taxonomic relatedness of *O. corvina* to other keratin-degrading fungi. Additionally, functional validation of the predicted proteases and CAZymes remains to be carried out.

Abbreviations

AA	Auxiliary activity enzymes
CAZymes	Carbohydrate-active enzymes
CE	Carbohydrate esterases
CBS	Centraalbureau voor Schimmelcultures (Central Bureau of Fungal Cultures)
EVM	Evidence Modeler
GH	Glycoside hydrolases
GH	Glycosyltransferases
Hotpep	Homology to Peptide Pattern pipeline
LPMOs	Lytic polysaccharide monooxygenases
N50	Assembly metric indicating contig length at 50% genome coverage
PE150	Paired-end 150 bp sequencing
PASA	Program to Assemble Spliced Alignments
SignalP	Signal peptide prediction software
TMHMM	Transmembrane helix prediction software
BUSCO	Benchmarking Universal Single-Copy Orthologs

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13104-025-07566-9.

Supplementary material 1. List of 38 putatively secreted carbohydrate-active enzymes encoded in the *Onygena corvina* genome. For each gene, the table provides the Gene ID, predicted CAZy family categorized as glycoside hydrolases (GH), glycosyltransferases (GT), carbohydrate esterases (CE), or auxiliary activity (AA) enzymes, annotated using dbCAN3, along with known enzymatic activities in each family according to the CAZy database.

Author contributions

S.P. and S.L.L.R. conceptualized the study, performed data analysis, prepared figures and drafted the manuscript. S.L.L.R. and V.E. critically revised and reviewed the manuscript.

Funding

This work was supported by the SFI Industrial Biotechnology program (project number 309558) funded by the Research Council of Norway.

Data availability
This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JBPRGC000000000. The version described in this paper is version JBPRGC010000000. The genome has been made available in the NCBI GenBank under the accession number PRJNA1280007. DNA and RNA raw reads are available in the Sequence Read Archive under the accession number PRJNA1327576. The mitochondrial genome has been deposited in the NCBI GenBank under the accession PX395520. The gene and protein annotations, putative functional annotation, and CAZyme predictions are available on Figshare (https://doi.org/10.6084/m9.figshare.29328893.v1).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 2 July 2025 / Accepted: 24 October 2025
Published online: 05 December 2025

References

1. Huang Y, Busk PK, Herbst F-A, Lange L. Genome and secretome analyses provide insights into keratin decomposition by novel proteases from the non-pathogenic fungus *Onygena corvina*. Appl Microbiol Biotechnol. 2015;99:9635–49. https://doi.org/10.1007/s00253-015-6805-9.
2. Lange L, Huang Y, Busk PK. Microbial decomposition of keratin in nature—a new hypothesis of industrial relevance. Appl Microbiol Biotechnol. 2016;100:2083–96. https://doi.org/10.1007/s00253-015-7262-1.
3. Huang Y, Busk PK, Lange L. Production and characterization of keratinolytic proteases produced by *Onygena corvina*. Fungal Genomics Biol. 2015;05:5–119. https://doi.org/10.4172/2165-8056.1000119.
4. Qiu J, Wilkens C, Barrett K, Meyer AS. Microbial enzymes catalyzing keratin degradation: Classification, structure, function. Biotechnol Adv. 2020;44:107607. https://doi.org/10.1016/j.biotechadv.2020.107607.
5. Kandemir H, Dukik K, de Melo Teixeira M, Stielow JB, Delma FZ, Al-Hatmi AMS, et al. Phylogenetic and ecological reevaluation of the order onygenales. Fungal Divers. 2022;115:1–72. https://doi.org/10.1007/s13225-022-00506-z.
6. May T, Vaughan L, Holmes G, Brand E, Pegrem-Brand F, Pegrem-Brand L, et al. Citizen scientists detect the fungus *Onygena corvina* (Onygenales, Ascomycota) in new South Wales, Australia. Aust J Taxon. 2024;72:1–11. https://doi.org/10.54102/ajt.i1ogn.
7. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20. https://doi.org/10.1093/bioinformatics/btu170.
8. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27:722–36. https://doi.org/10.1101/gr.215087.116.
9. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 2017;27:737–46. https://doi.org/10.1101/gr.214270.116.
10. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE. 2014;9:e112963. https://doi.org/10.1371/journal.pone.0112963.
11. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinforma Oxf Engl. 2015;31:3210–2. https://doi.org/10.1093/bioinformatics/btv351.
12. Xu Z, Wang H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. 2007;35 suppl2:W265–8. https://doi.org/10.1093/nar/gkm286.
13. Han Y, Wessler SR. MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. Nucleic Acids Res. 2010;38:e199. https://doi.org/10.1093/nar/gkq862.
14. Price AL, Jones NC, Pevzner PA. De Novo identification of repeat families in large genomes. Bioinforma Oxf Engl. 2005;21(Suppl 1):i351–358. https://doi.org/10.1093/bioinformatics/bti1018.
15. Edgar RC, Myers EW. PILER: Identification and classification of genomic repeats. Bioinformatics. 2005;21 suppl_1:i152–8. https://doi.org/10.1093/bioinformatics/bti1003
16. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 2007;8:973–82. https://doi.org/10.1038/nrg2165.
17. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res. 2005;110:462–7. https://doi.org/10.1159/000084979.
18. Chen N. Using repeatmasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinforma Ed Board Andreas Baxevanis Al. 2004;Chap 4(Unit 410). https://doi.org/10.1002/0471250953.bi0410s05.
19. Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA. J Mol Biol. 1997;268:78–94. https://doi.org/10.1006/jmbi.1997.0951.
20. Majoros WH, Pertea M, Salzberg SL. TigrScan and glimmerhmm: two open source Ab initio eukaryotic gene-finders. Bioinformatics. 2004;20:2878–9. https://doi.org/10.1093/bioinformatics/bth315.
21. Blanco E, Parra G, Guigó R. Using Geneid to identify genes. Curr Protoc Bioinforma. 2007;Chap4–Unit43. https://doi.org/10.1002/0471250953.bi0403s18.
22. Korf I. Gene finding in novel genomes. BMC Bioinformatics. 2004;5:59. https://doi.org/10.1186/1471-2105-5-59.
23. Stanke M, Waack S. Gene prediction with a hidden Markov model and a new intron submodel. Bioinforma Oxf Engl. 2003;19(Suppl 2):ii215–225. https://doi.org/10.1093/bioinformatics/btg1080.
24. Keilwagen J, Wenk M, Erickson JL, Schattat MH, Grau J, Hartung F. Using intron position conservation for homology-based gene prediction. Nucleic Acids Res. 2016;44:e89. https://doi.org/10.1093/nar/gkw092.
25. Pertea M, Kim D, Pertea G, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie, and ballgown. Nat Protoc. 2016;11:1650–67. https://doi.org/10.1038/nprot.2016.095.
26. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De Novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. Nat Protoc. 2013;8:1494–512. https://doi.org/10.1038/nprot.2013.084.
27. Campbell MA, Haas BJ, Hamilton JP, Mount SM, Buell CR. Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. BMC Genomics. 2006;7:327. https://doi.org/10.1186/1471-2164-7-327.
28. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, et al. Automated eukaryotic gene structure annotation using EVidence Modeler and the program to assemble spliced alignments. Genome Biol. 2008;9:R7. https://doi.org/10.1186/gb-2008-9-1-r7.
29. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34:3094–100. https://doi.org/10.1093/bioinformatics/bty191.
30. Cabanettes F, Klopp C. D-GENIES: Dot plot large genomes in an interactive, efficient and simple way. PeerJ. 2018;6:e4958. https://doi.org/10.7717/peerj.4958.
31. Blin K, Shaw S, Vader L, Szenei J, Reitz ZL, Augustijn HE, et al. AntiSMASH 8.0: Extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. Nucleic Acids Res. 2025;53(W1):W32–8. https://doi.org/10.1093/nar/gkaf334.
32. Tamano K, Kuninaga M, Kojima N, Umemura M, Machida M, Koike H. Use of the KojA promoter, involved in Kojic acid biosynthesis, for polyketide production in *Aspergillus oryzae*: implications for long-term production. BMC Biotechnol. 2019;19:70. https://doi.org/10.1186/s12896-019-0567-x.
33. Slightom JL, Metzger BP, Luu HT, Elhammer AP. Cloning and molecular characterization of the gene encoding the Aureobasidin A biosynthesis complex in aureobasidium pullulans BP-1938. Gene. 2009;431:67–79. https://doi.org/10.1016/j.gene.2008.11.011.
34. Wollenberg RD, Saei W, Westphal KR, Klitgaard CS, Nielsen KL, Lysøe E, et al. Chrysogine biosynthesis is mediated by a two-module nonribosomal peptide synthetase. J Nat Prod. 2017;80:2131–5. https://doi.org/10.1021/acs.jnatprod.6b00822.
35. Godio RP, Fouces R, Martín JF. A squalene epoxidase is involved in biosynthesis of both the antitumor compound clavaric acid and sterols in the basidiomycete *H. sublateritium*. Chem Biol. 2007;14:1334–46. https://doi.org/10.1016/j.chembiol.2007.10.018.

36. Zhang W-Y, Zhong Y, Yu Y, Shi D-F, Huang H-Y, Tang X-L, et al. 4-Hydroxy pyridones from heterologous expression and cultivation of the native host. J Nat Prod. 2020;83:3338–46. https://doi.org/10.1021/acs.jnatprod.0c00675.
37. Rawlings ND, Barrett AJ, Bateman A. MEROPS: the peptidase database. Nucleic Acids Res. 2010;38 suppl_1:D227–33. https://doi.org/10.1093/nar/gkp971
38. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019;37:420–3. https://doi.org/10.1038/s41587-019-0036-z.
39. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes1. J Mol Biol. 2001;305:567–80. https://doi.org/10.1006/jmbi.2000.4315.
40. Zheng J, Ge Q, Yan Y, Zhang X, Huang L, Yin Y. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. Nucleic Acids Res. 2023;51:W115–21. https://doi.org/10.1093/nar/gkad328.
41. Busk PK, Lange L. Classification of fungal and bacterial lytic polysaccharide monooxygenases. BMC Genomics. 2015;16:368. https://doi.org/10.1186/s12864-015-1601-6.
42. Hemsworth GR, Henrissat B, Davies GJ, Walton PH. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. Nat Chem Biol. 2014;10:122–6. https://doi.org/10.1038/nchembio.1417.
43. Støpamo FG, Røhr ÅK, Mekasha S, Petrović DM, Várnai A, Eijsink VGH. Characterization of a lytic polysaccharide monooxygenase from *Aspergillus fumigatus* shows functional variation among family AA11 fungal LPMOs. J Biol Chem. 2021;297:101421. https://doi.org/10.1016/j.jbc.2021.101421.
44. Rieder L, Petrović D, Väljamäe P, Eijsink VGH, Sørlie M. Kinetic characterization of a putatively Chitin-Active LPMO reveals a preference for soluble substrates and absence of monooxygenase activity. ACS Catal. 2021;11:11685–95. https://doi.org/10.1021/acscatal.1c03344.

Publisher's note