Syracuse University
College of Arts & Sciences
Forensic & National
Security Sciences Institute

ATCC®

Yarrowia lipolytica: A Versatile Non-conventional Oleaginous Yeast for Biotechnology

James E. Crill
Professor of Practice, Syracuse University

Scott V. Nguyen, PhD
Senior Biocuration Scientist, ATCC

Welcome & Introductions

Syracuse University
College of Arts & Sciences
Forensic & National Security Sciences Institute

ATCC®

Prof. James E. Crill

Professor of Practice, Syracuse University

James teaches microbial forensics and biochemical analysis while conducting research on *Yarrowia lipolytica* genomics and biomanufacturing. His work focuses on comparative genomics, strain engineering, and forensic applications of microbial systems. Previously, James led deployable microbial forensic capabilities at MRIGlobal, developing sequencing and detection methods for biological threat identification.

Scott V. Nguyen, PhD

Sr. Biocuration Scientist, ATCC

Scott has described several novel microbial species and identified multiple new SARS-CoV-2 variants including the Delta-Omicron recombinant XD variant popularly known as 'Deltacron'. He previously held positions at the USDA-ARS, University College Dublin, and at the DC Public Health Laboratory. He has a PhD in microbiology and immunology from the University of Oklahoma Health Sciences Center.

About ATCC

ATCC is a global leader in providing authenticated, high-quality biological resources and standards for industry, academia, and government.

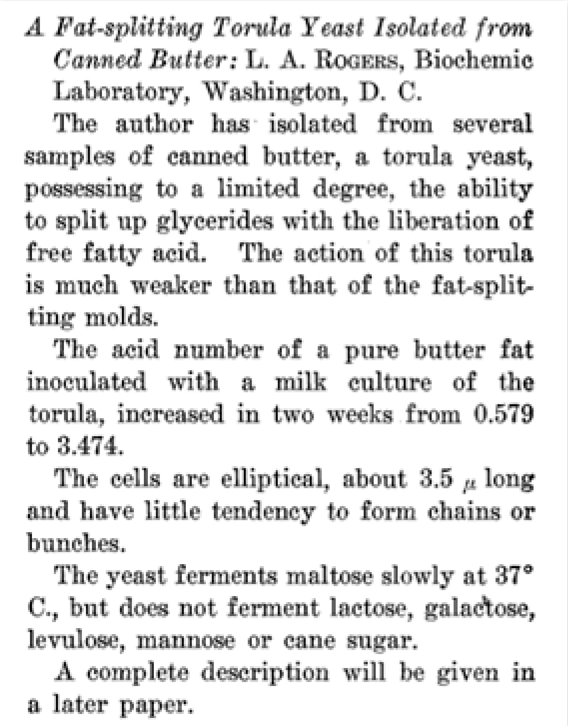
- Founded in 1925, ATCC is a private, nonprofit, global biological resource center and standards organization that provides scientists with the biomaterials and resources they need to conduct critical life science research.
- Worldwide biological materials resource and standards development organization:
    - ATCC Genome Portal
        - 6,500 genomes
        - 3,000 transcriptomes
        - 500 exomes
    - 80,000+ microorganisms
    - 4,000+ cell lines
    - Genomic & synthetic nucleic acids
    - Standards & controls
    - Media, sera, and reagents
    - Advanced cell models

www.atcc.org/

Yarrowia lipolytica

Historical significance and early industrial relevance

- *Y. lipolytica* is a non-conventional oleaginous yeast that was first identified in the early 20th century as a spoiling agent of butter and margarine
- Lore A. Rogers (co-founder of ATCC) characterized *Y. lipolytica* in his early research on butter spoilage organisms
- Historically known under various synonyms
- Control of *Y. lipolytica* linked to stability of margarine (Jacobsen of Anton Jurgens' Margarinefabrieken N.V., now known as Unilever)



A Fat-splitting Torula Yeast Isolated from Canned Butter: L. A. ROGERS, Biochemic Laboratory, Washington, D. C.

The author has isolated from several samples of canned butter, a torula yeast, possessing to a limited degree, the ability to split up glycerides with the liberation of free fatty acid. The action of this torula is much weaker than that of the fat-splitting molds.

The acid number of a pure butter fat inoculated with a milk culture of the torula, increased in two weeks from 0.579 to 3.474.

The cells are elliptical, about 3.5 μ long and have little tendency to form chains or bunches.

The yeast ferments maltose slowly at 37° C., but does not ferment lactose, galactose, levulose, mannose or cane sugar.

A complete description will be given in a later paper.

Rogers LA. 1904. Studies upon the keeping quality of butter. I. Canned butter. Govt. print. off, Washington, 1904. Accessed Jan. 12, 2026. https://www.biodiversitylibrary.org/bibliography/48941.

Duality of *Yarrowia*

From contaminant to biotechnology applications

Spoilage & contaminant
Lipolytic activity leads to off-flavors in dairy and lipid-rich fermentations.

Environmental resilience
Tolerates weak acids, salt, oxidative stress, low pH, and hydrophobic media.

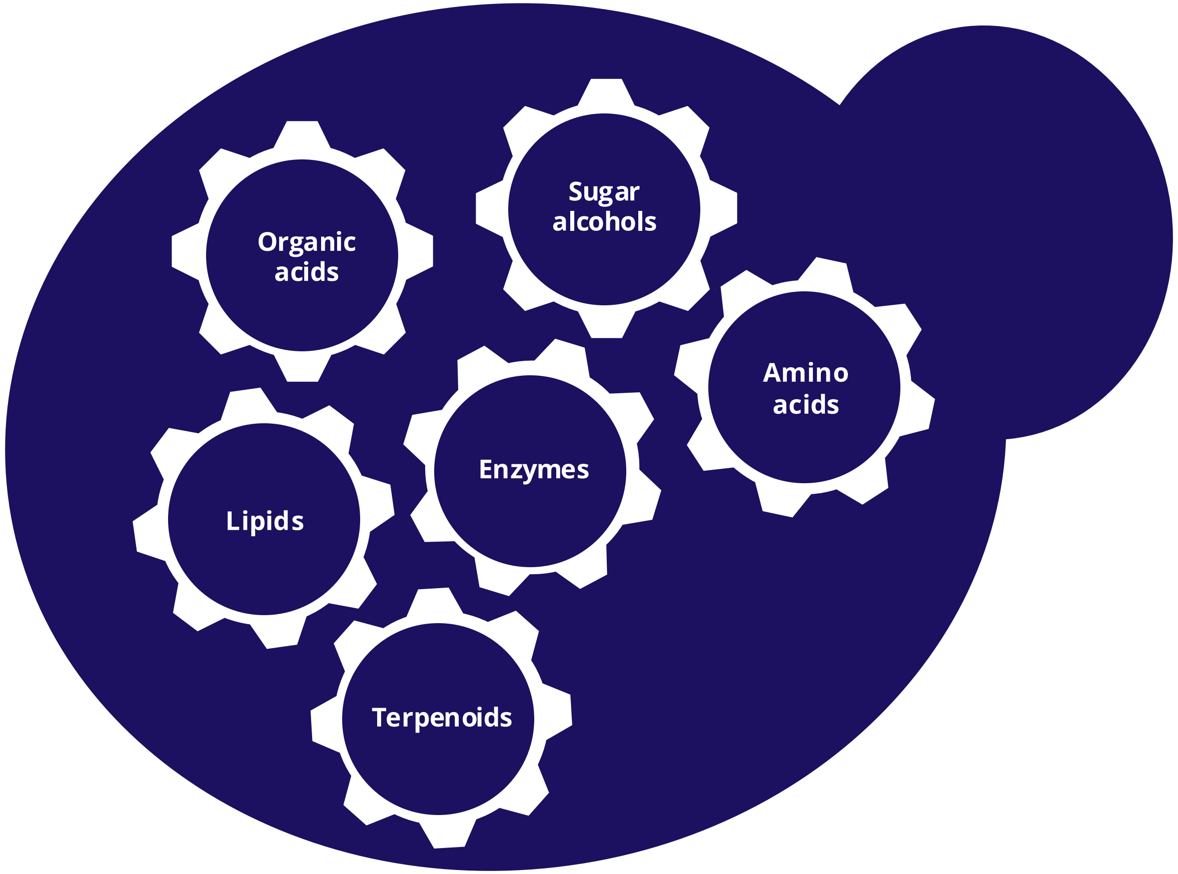
Biotechnological advantage
Traits enabling survival in harsh environments support waste valorization.

Why *Yarrowia lipolytica*?

Syracuse University
College of Arts & Sciences
Forensic & National
Security Sciences Institute

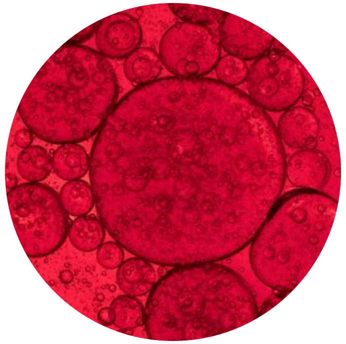
ATCC®

- **Highly oleaginous:** ~30% lipid content in wild-type strains; >50% in engineered strains
- **Industrial relevance:** Useful for biofuel production, oleochemical synthesis, waste valorization, and protein expression
- **Superior lipid metabolism:** Outperforms *S. cerevisiae* in lipid metabolism and utilization of non-conventional substrates



Ecological Niches of *Y. lipolytica*

Widely distributed in soil, seawater, dairy, and lipid-rich environments



Specializes in degradation of alkanes, fatty acids, oils, and fats

Secretes enzymes that reshape local microbial communities

Plays a role in carbon cycling and bioremediation

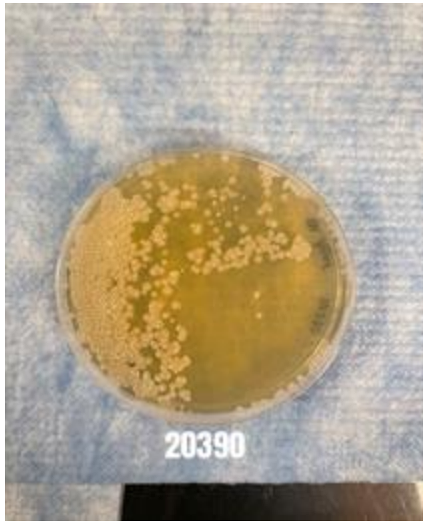
Phenotypic Diversity in *Yarrowia*

Morphology, metabolism, and strain-level specialization

- Wide phenotypic variation across strains in morphology and metabolism
- Variations in lipid accumulation capacity, fatty acid profiles, substrate utilization ranges, and stress tolerance
- Diverse morphologies and metabolic capabilities suggest strain-level specialization within the species

Dry/White

Dry/White



Oily/Tan

Very Oily/Opaque

Early Biotechnical Applications of *Y. lipolytica*

Early biotech applications

- Pioneered in the 1950s by British Petroleum (BP) for industrial biotechnology
- Explored as a source of single-cell protein (SCP)
- Utilized n-alkanes (hydrocarbons) as carbon sources
- Commercial product Toprina G developed as protein-rich animal feed

Significance of single-celled proteins (SCP)

- Microbial protein produced from non-traditional substrates
- Sustainable alternative to conventional protein sources
- European Food Safety Authority (EFSA) concluded *Y. lipolytica* biomass is safe and nutritionally adequate

Science Museum Group. Specimen of "Toprina" single cell protein. 1977-438 Science Museum Group Collection Online. Accessed 15 August 2025. https://collection.sciencemuseumgroup.org.uk/objects/co61213/specimen-of-toprina-single-cell-protein.

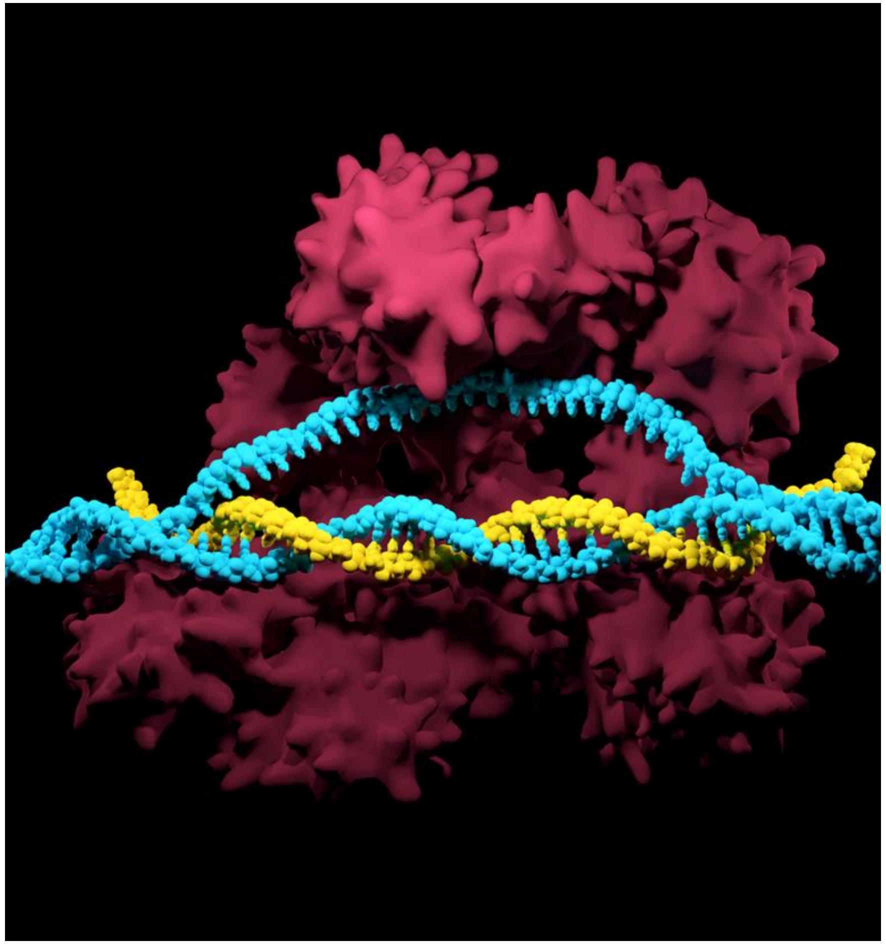
Current State of Research

- Mature genetic toolbox
    - CRISPR/Cas9, Golden Gate assembly, synthetic biology platforms
- Industrial readiness
    - FDA GRAS status
    - Proven commercial-scale fermentation
- Primary applications
    - Single-cell oils, citric acid, recombinant proteins, biosurfactants
- Feedstock flexibility
    - Waste oils, glycerol, lignocellulosic biomass
- Well-characterized metabolism
    - Lipid biosynthesis and organic acid pathways

Genetic Engineering Tools

Molecular tools for *Yarrowia*

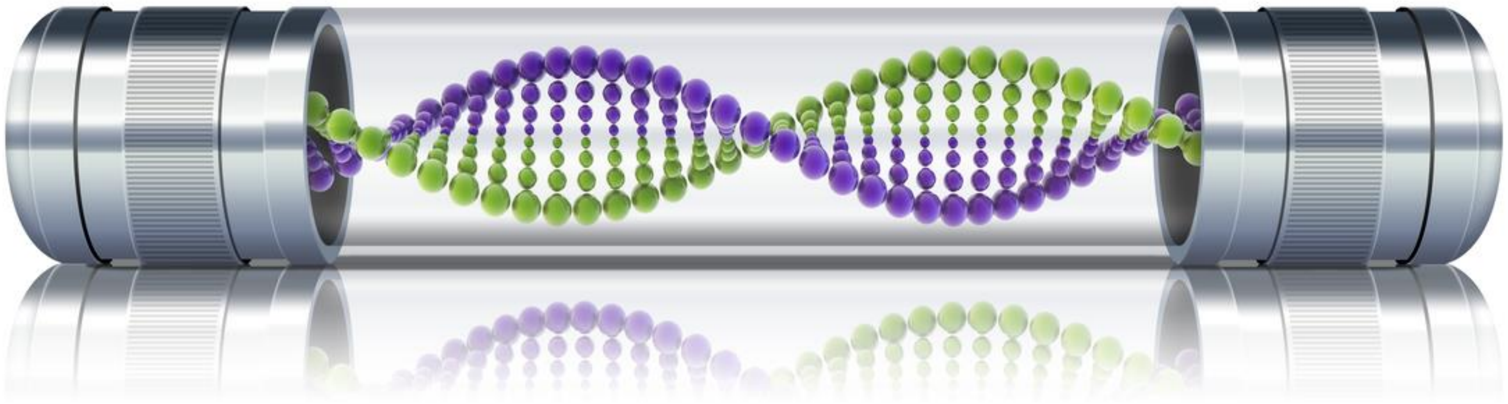
- Early tools (1990s):
  - URA3 blaster and Cre/loxP for marker recycling
  - Heterologous gene integration
- Programmable nucleases
  - TALENs for targeted knock-ins/knock-outs
  - Limited by efficiency and labor intensity
- CRISPR/Cas systems (≈2016–present)
  - CRISPR/Cas9 for efficient genome editing
  - Multiplex editing (e.g., TRP1, PEX10, GUT2)
  - CRISPRi/a for gene repression and activation
- Recent innovations include:
  - Promoter swapping for fine-tuned gene expression
  - Genome-wide CRISPR libraries for functional screening



Synthetic Biology Tools

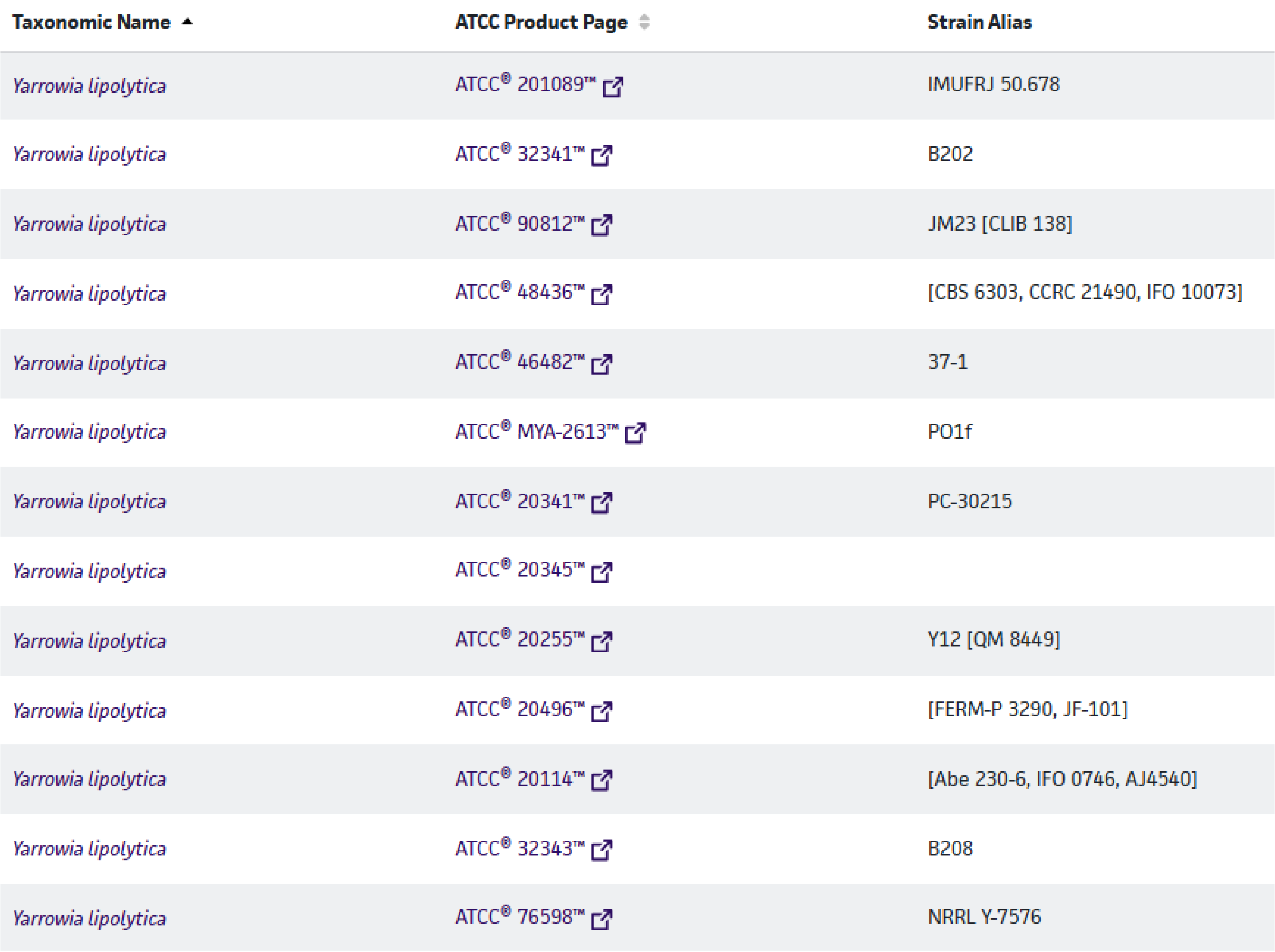
Efficiency, challenges, and optimization strategies

- Strain domestication
  - Auxotrophic markers (*URA3*, *LEU2*, *LYS5*)
  - Protease deletions (*AEP*, *AXP*) for protein stability
  - Morphology control via *MHY1* deletion
- Regulatory systems
  - Inducible promoters (e.g., erythritol-inducible systems)
- Chassis Development
  - Enhanced precursor supply (acetyl-CoA, NADPH)
  - Pathway optimization for lipids, terpenoids, aromatics
- Precision control
  - CRISPR/Cas9 and CRISPRi/a for genome-level regulation



Yarrowia Collection at ATCC

https://genomes.atcc.org/



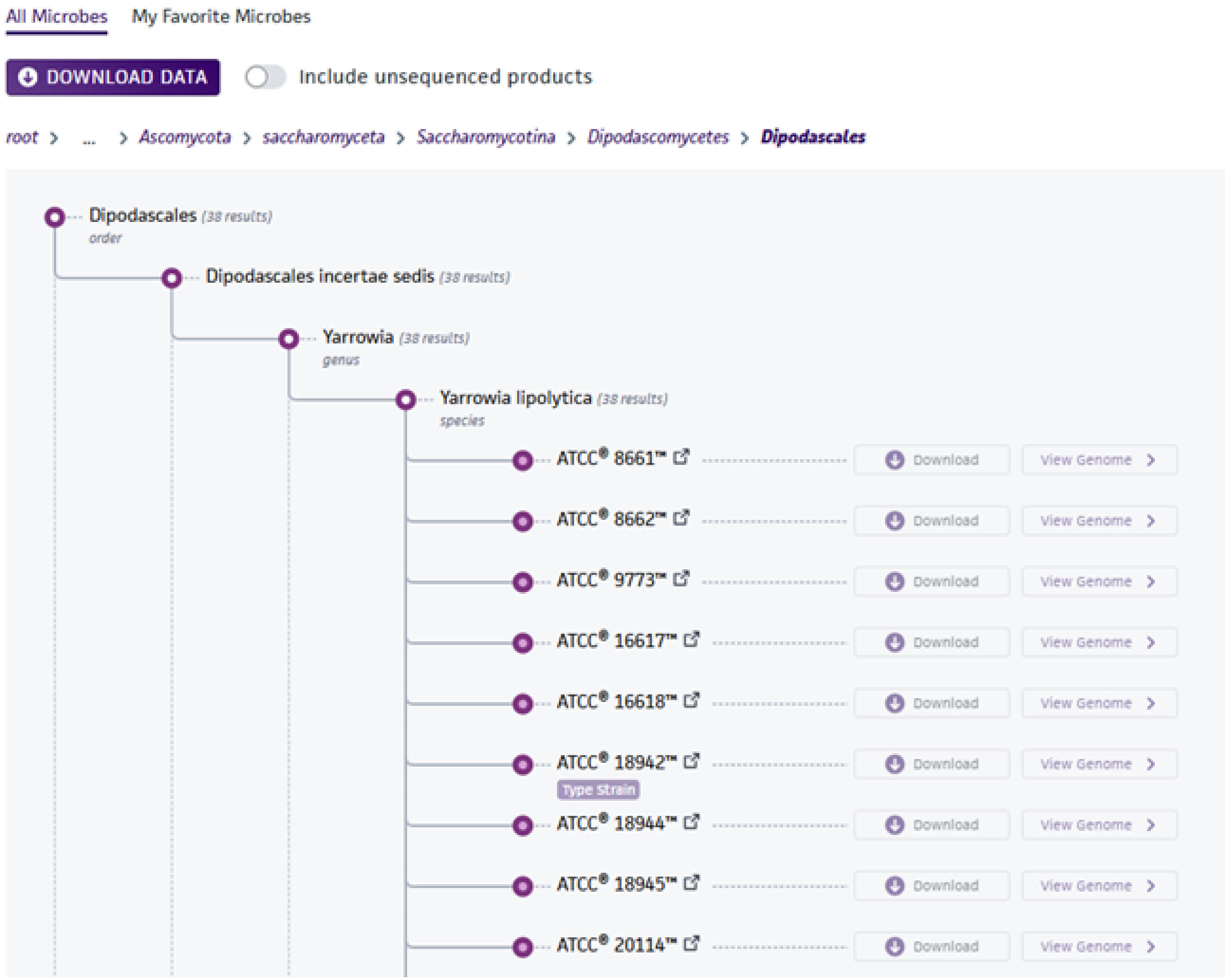
Taxonomic Name ▲	ATCC Product Page ⇕	Strain Alias
Yarrowia lipolytica	ATCC® 201089™	IMUFRJ 50.678
Yarrowia lipolytica	ATCC® 32341™	B202
Yarrowia lipolytica	ATCC® 90812™	JM23 [CLIB 138]
Yarrowia lipolytica	ATCC® 48436™	[CBS 6303, CCRC 21490, IFO 10073]
Yarrowia lipolytica	ATCC® 46482™	37-1
Yarrowia lipolytica	ATCC® MYA-2613™	PO1f
Yarrowia lipolytica	ATCC® 20341™	PC-30215
Yarrowia lipolytica	ATCC® 20345™	
Yarrowia lipolytica	ATCC® 20255™	Y12 [QM 8449]
Yarrowia lipolytica	ATCC® 20496™	[FERM-P 3290, JF-101]
Yarrowia lipolytica	ATCC® 20114™	[Abe 230-6, IFO 0746, AJ4540]
Yarrowia lipolytica	ATCC® 32343™	B208
Yarrowia lipolytica	ATCC® 76598™	NRRL Y-7576

ATCC Genome Portal (“AGP”)

- 38 *Yarrowia* genomes
  - Sequenced through hybrid approach (short reads + long reads)
  - Many of the strains are used in biotech
- Fully annotated
  - Only 14 *Yarrowia* genomes in NCBI are annotated

Yarrowia Collection at ATCC

https://genomes.atcc.org/



root > ... > *Ascomycota* > *saccharomyceta* > *Saccharomycotina* > *Dipodascomycetes* > ***Dipodascales***

- Dipodascales *(38 results)* *order*
  - Dipodascales incertae sedis *(38 results)*
    - Yarrowia *(38 results)* *genus*
      - Yarrowia lipolytica *(38 results)* *species*
        - ATCC® 8661™
        - ATCC® 8662™
        - ATCC® 9773™
        - ATCC® 16617™
        - ATCC® 16618™
        - ATCC® 18942™ Type Strain
        - ATCC® 18944™
        - ATCC® 18945™
        - ATCC® 20114™

ATCC Genome Portal (“AGP”)
NEW Taxonomy Tree

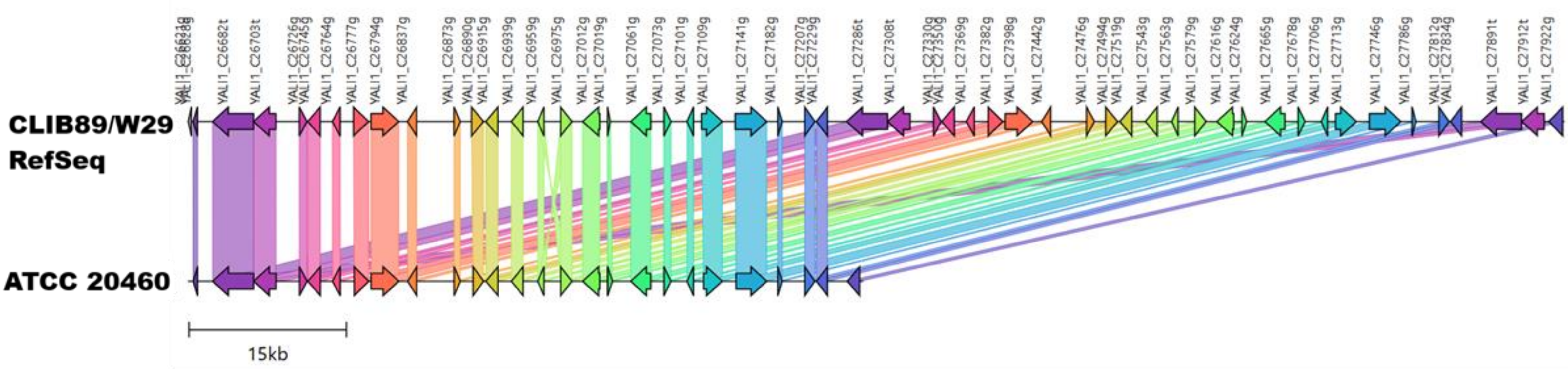
- Easily search the ATCC Genome Portal with the taxonomy tree
- Interactive taxonomic navigation
- Advanced search & filtering
- Integrated data access

Yarrowia Collection at ATCC

Superior assembly of ATCC® 20460™ (=W29 = CLIB89)

A 54 kb repeat is present in the CLIB89/W29 NCBI RefSeq assembly

- This 54 kb repeat is not observed in ATCC® 20460™ (=W29 = CLIB89)
- The repeat in the RefSeq assembly may have resulted from misassembly or a laboratory adapted variant of W29

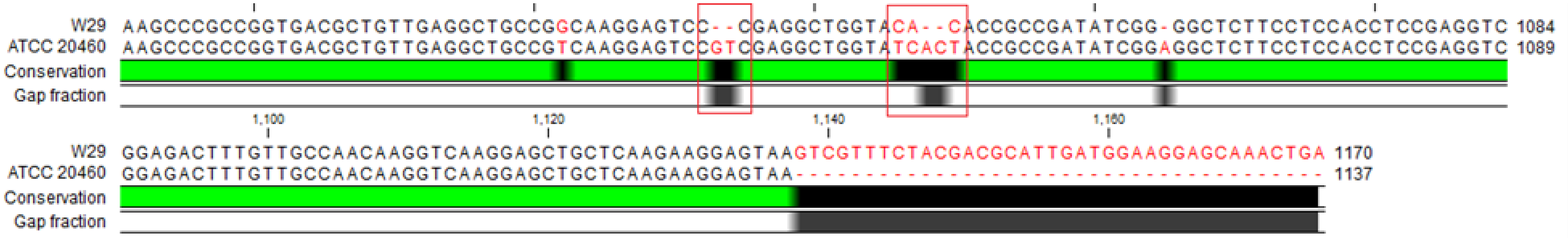


Yarrowia Collection at ATCC

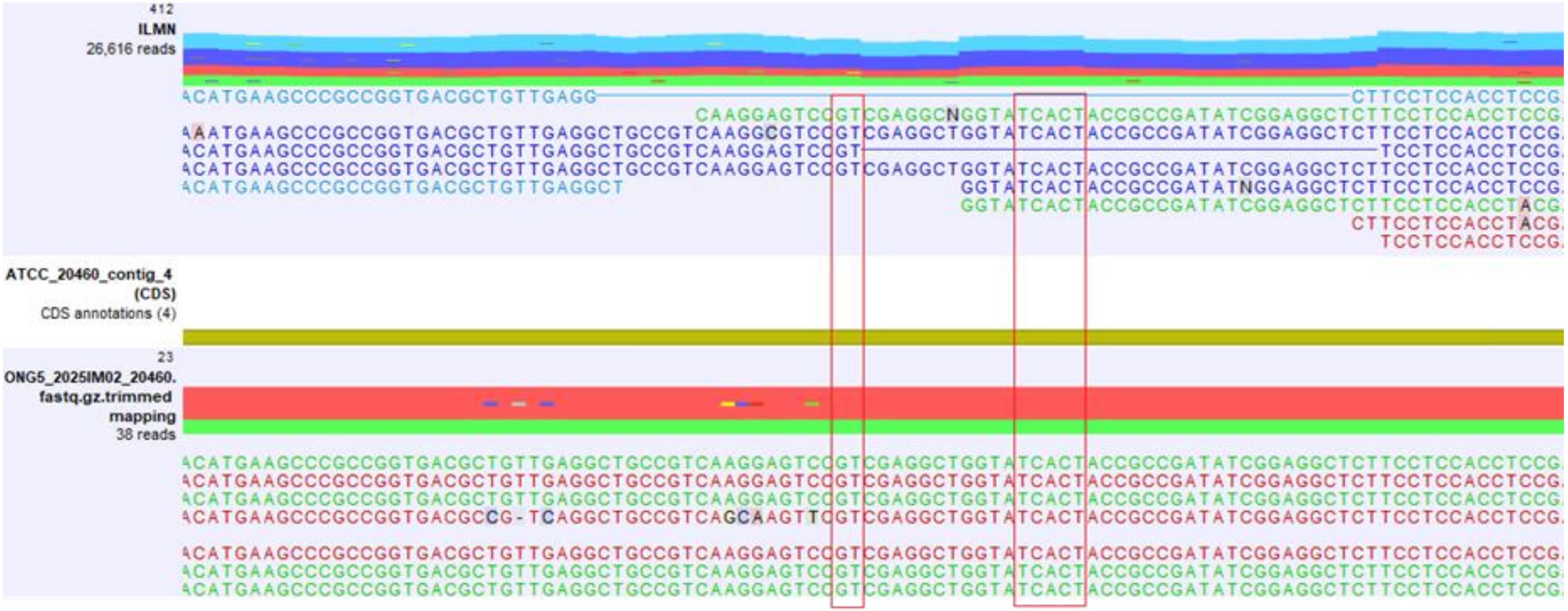
Correctly assembled *leu2* gene in the ATCC Genome Portal (“AGP”)

Syracuse University
College of Arts & Sciences
Forensic & National
Security Sciences Institute

ATCC®



```
W29         AAGCCCGCCGGTGACGCTGTTGAGGCTGCCGGCAAGGAGTCC--CGAGGCTGGTACA--CACCGCCGATATCGG-GGCTCTTCCTCCACCTCCGAGGTC 1084
ATCC 20460  AAGCCCGCCGGTGACGCTGTTGAGGCTGCCGTCAAGGAGTCCGTCGAGGCTGGTATCACTACCGCCGATATCGGAGGCTCTTCCTCCACCTCCGAGGTC 1089
Conservation
Gap fraction

                   1,100                1,120                1,140                1,160
W29         GGAGACTTTGTTGCCAACAAGGTCAAGGAGCTGCTCAAGAAGGAGTAAGTCGTTTCTACGACGCATTGATGGAAGGAGCAAACTGA 1170
ATCC 20460  GGAGACTTTGTTGCCAACAAGGTCAAGGAGCTGCTCAAGAAGGAGTAA-------------------------------------- 1137
Conservation
Gap fraction
```



Yarrowia Collection at ATCC

Correctly assembled *leu2* gene in the ATCC Genome Portal (“AGP”)

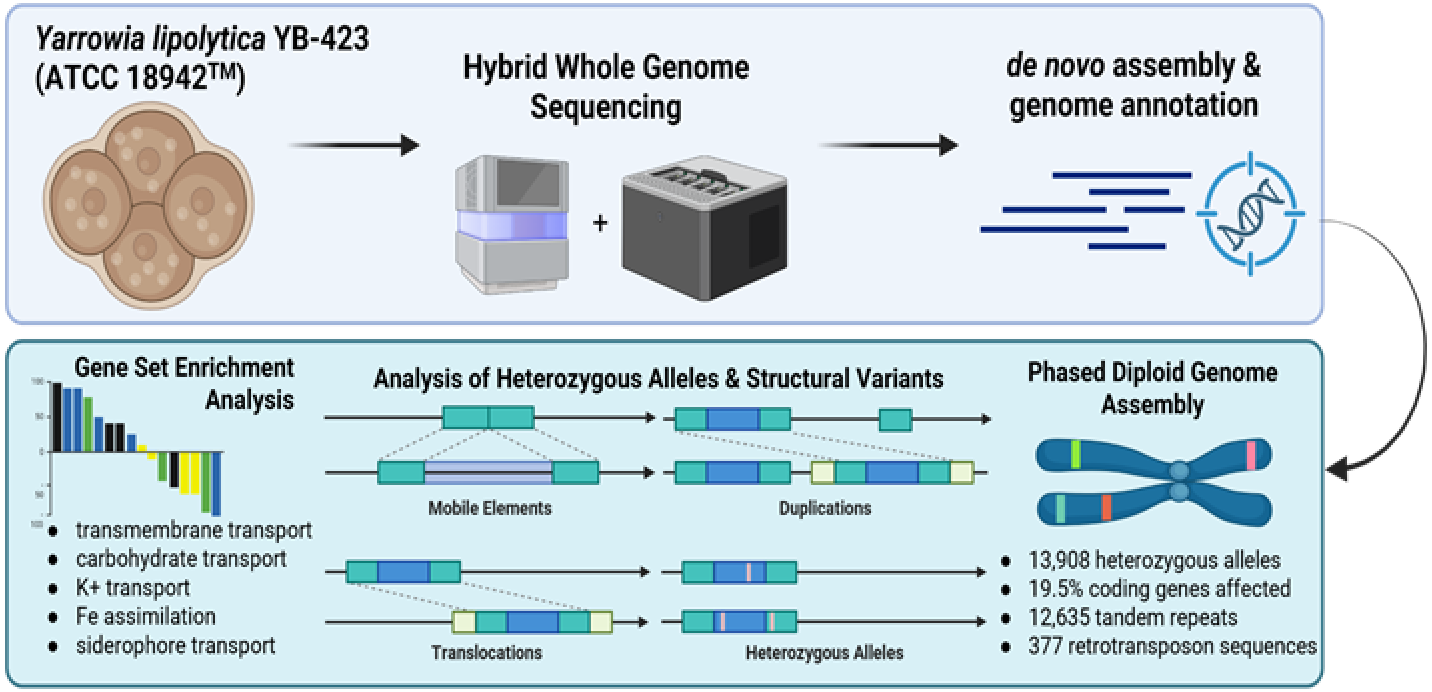
- Indels in the NCBI RefSeq assembly (GCF_001761485.1) result in C-terminal variations of LEU2
- LEU2 in ATCC® 20460™ matches other LEU2 annotations from other *Y. lipolytica* in the ATCC Genome Portal and other NCBI assemblies

```
                                   20        40        60        80
W29 leu2          MEPETKKTKTDSKKIVLLGGDFCGPEVIAEAVKVLKSVAEASGTEFVFEDRLIGGAAIEKEGEPITDATLDICRKADSIMLGAVGGAANTVWTTPDGR 98
ATCC 20460 leu2   MEPETKKTKTDSKKIVLLGGDFCGPEVIAEAVKVLKSVAEASGTEFVFEDRLIGGAAIEKEGEPITDATLDICRKADSIMLGAVGGAANTVWTTPDGR 98
Conservation
Gap fraction

                 100       120       140       160       180
W29 leu2          TDVRPEQGLLKLRKDLNLYANLRPCQLLSPKLADLSPIRNVEGTDFIIVRELVGGIYFGERKEDDGSGVASDTETYSVPEVERIARMAAFLALQHNPP 196
ATCC 20460 leu2   TDVRPEQGLLKLRKDLNLYANLRPCQLLSPKLADLSPIRNVEGTDFIIVRELVGGIYFGERKEDDGSGVASDTETYSVPEVERIARMAAFLALQHNPP 196
Conservation
Gap fraction

                    200       220       240       260       280
W29 leu2          LPVWSLDKANVLASSRLWRKTVTRVLKDEFPQLELNHQLIDSAAMILIKQPSKMNGIIITTNMFGDIISDEASVIPGSLGLLPSASLASLPDTNEAFG 294
ATCC 20460 leu2   LPVWSLDKANVLASSRLWRKTVTRVLKDEFPQLELNHQLIDSAAMILIKQPSKMNGIIITTNMFGDIISDEASVIPGSLGLLPSASLASLPDTNEAFG 294
Conservation
Gap fraction

                      300       320       340       360       380
W29 leu2          LYEPCHGSAPDLGKQKVNPIATILSAAMMLKFSLNMKPAGDAVEAAGKES--RGWYTPPISGL--------FLHLRGRRLCCQQGQGAAQEGVSRFYD 382
ATCC 20460 leu2   LYEPCHGSAPDLGKQKVNPIATILSAAMMLKFSLNMKPAGDAVEAAVKESVEAGITTADIGGSSSTSEVGDFVANKVKELLKKE*------------- 379
Conservation
Gap fraction

                        400
W29 leu2          ALMEGAN* 390
ATCC 20460 leu2   -------- 379
Conservation
Gap fraction
```

Type Strain Project

ATCC® 18942™

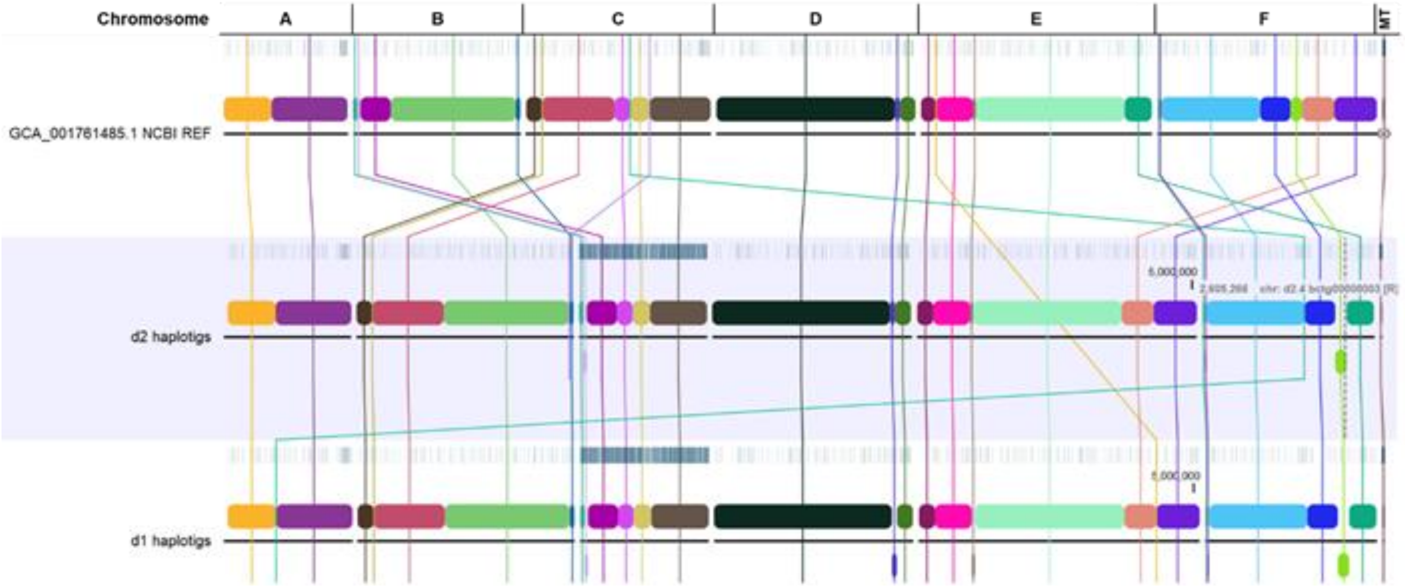
- First high-quality phased diploid assembly of the type strain ATCC® 18942™
- Provides a definitive taxonomic and genomic reference
- Captures heterozygosity and structural variation absent from haploid genomes
- Enables improved metabolic modeling, strain engineering, and forensic tracking



Crill JE II, et al. Complete Diploid Genome of the Type Strain *Yarrowia lipolytica* YB-423 (ATCC 18942™). J Ind Microbiol Biotechnol 8: kuag002, 2026. PubMed: 41504366
Diagram partially created with BioRender.com

Whole-genome Alignment

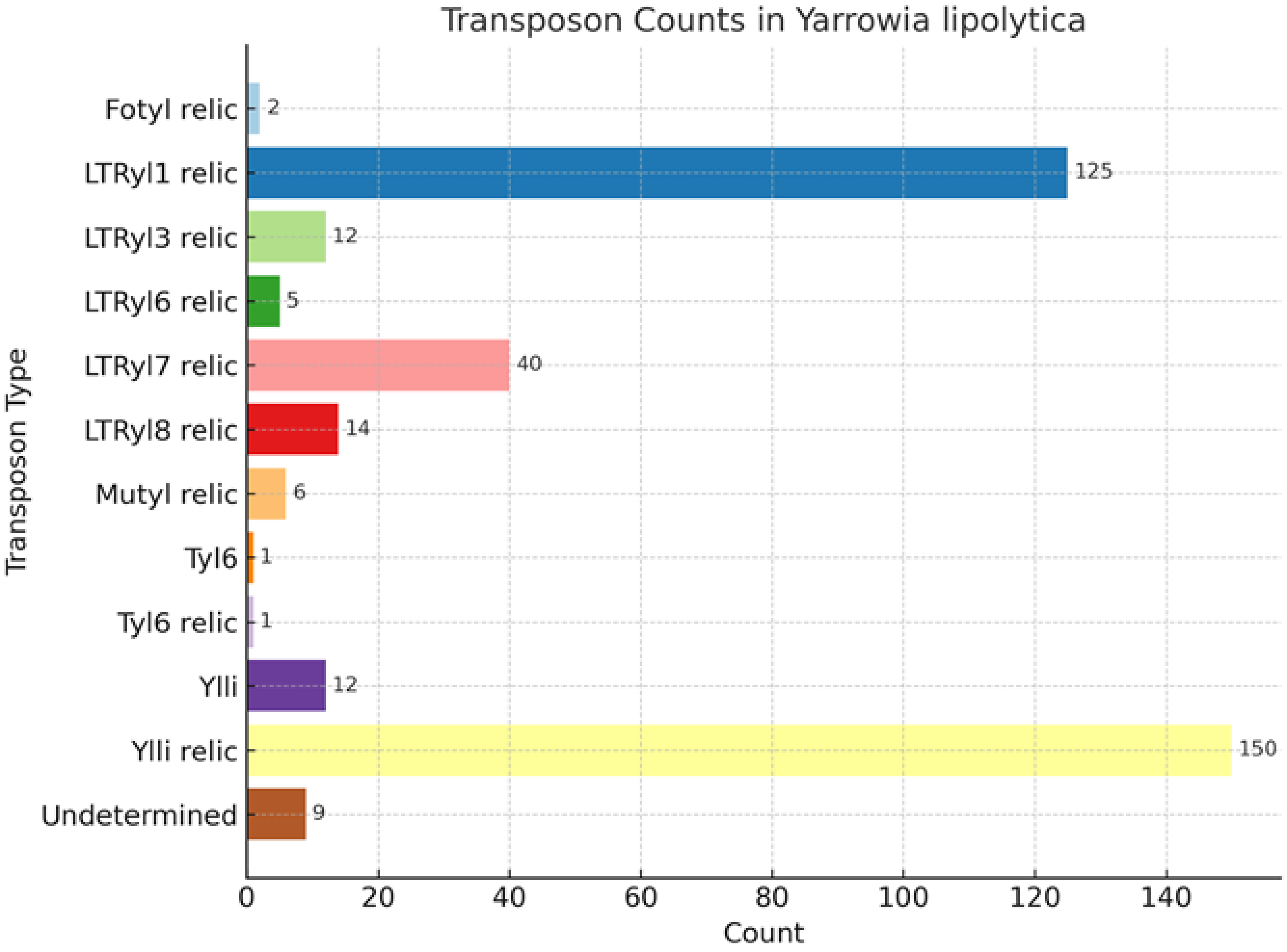
ATCC® 18942™ vs. W29



- Comparison of type strain ATCC® 18942™ and reference W29 reveals extensive chromosomal rearrangements
- Multiple reciprocal terminal translocations (RTTs) detected, with the largest spanning ~1.36 Mb
- RTT breakpoints lack transposons or mobile elements, but the rearrangement mechanisms remain unclear
- High genomic plasticity at structural level despite >99.4% sequence similarity across the species

Transposon Composition in the Type Strain

- Diverse transposable element families present, but mostly inactive remnants of past activity
- Low current TE mobility indicates genomically stable species with occasional large structural changes
- TE fingerprints offer potential for strain identification and forensic applications
- Historical TE bursts likely contributed to the chromosomal rearrangements observed between strains

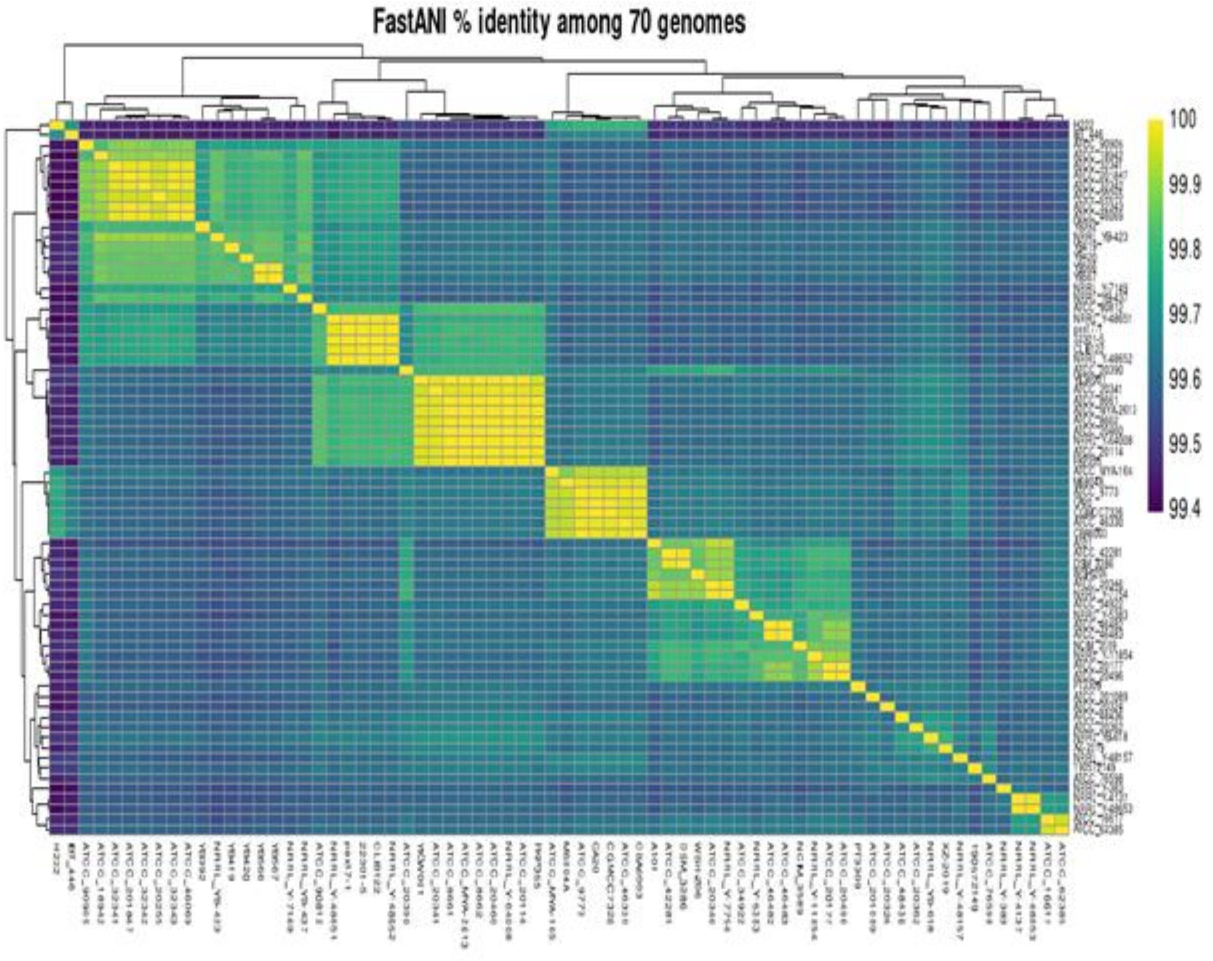


Average Nucleotide Analysis

Syracuse University College of Arts & Sciences
Forensic & National Security Sciences Institute

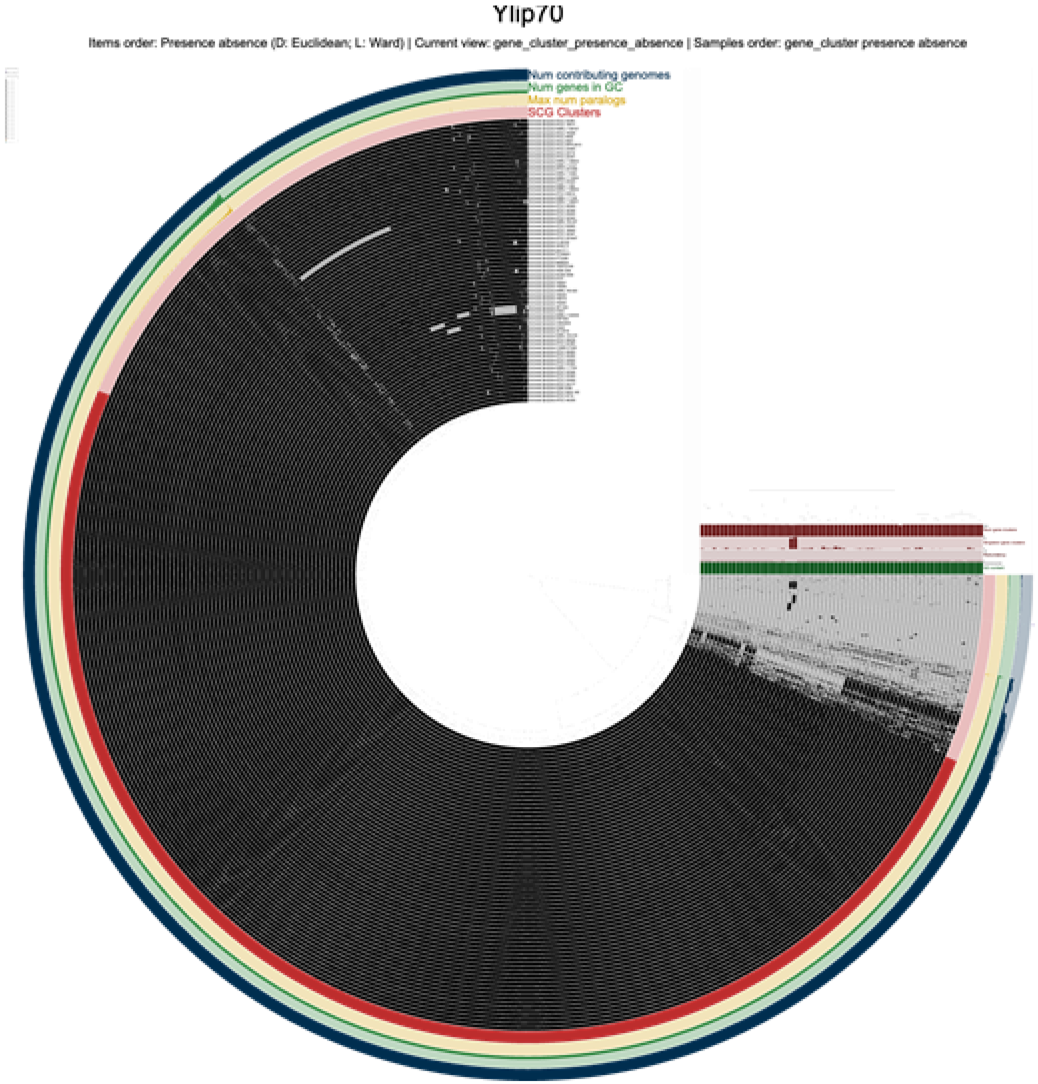
ATCC®

- All 70 strains demonstrate high sequence conservation
- Multiple distinct phylogenetic clades
- No deep divergence detected all strains fall within tight species boundaries despite a worldwide collection
- Hierarchical clustering reveals structured population with clear evolutionary lineages



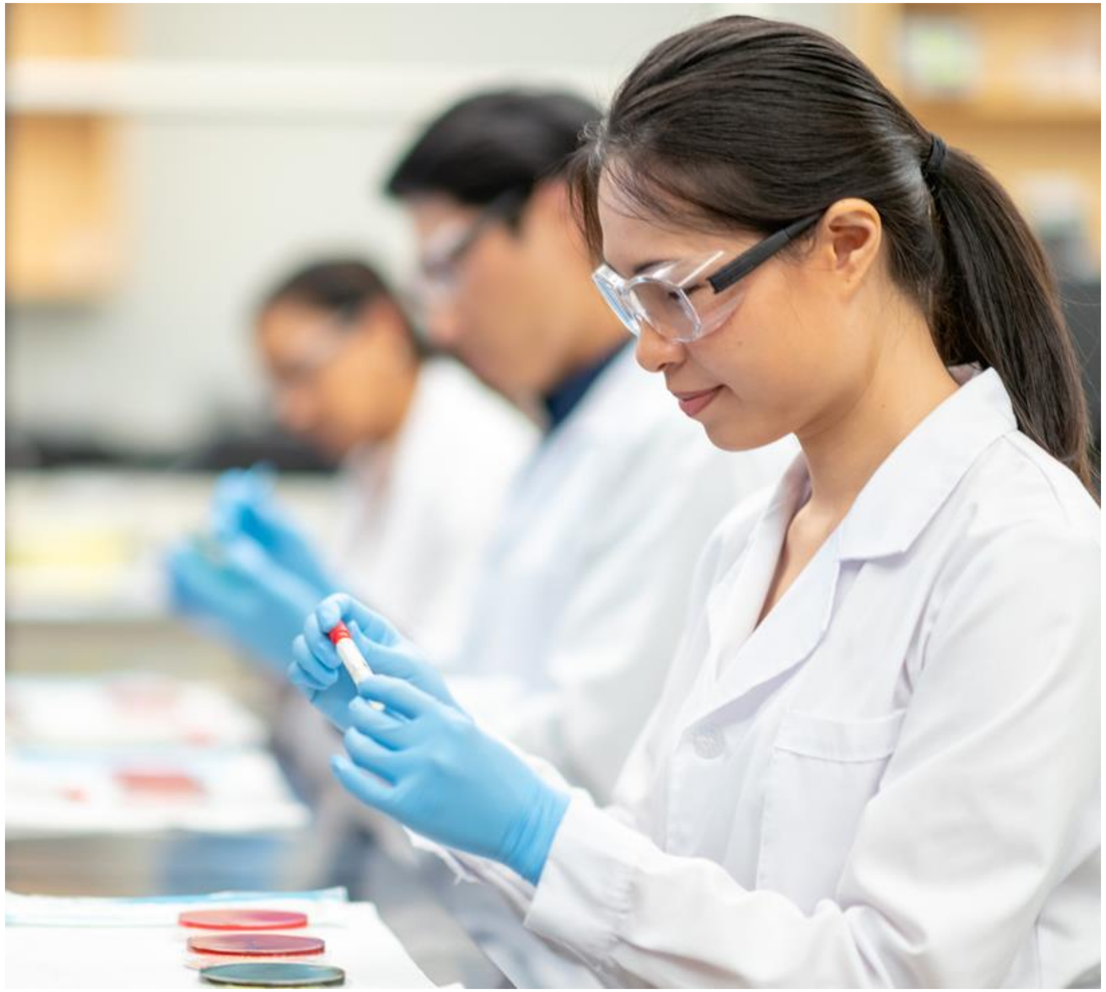
Pangenome Analysis

- Large core genome
- Limited accessory genome that indicates closed or moderately closed pangenome architecture
- Accessory genes show phylogenetically structured distribution, not random variation
- Gene presence/absence patterns correlate with strain relationships, indicating vertical inheritance



Conclusions

- Exceptional metabolic flexibility and lipid accumulation underpin the industrial value of *Yarrowia lipolytica*
- Despite high sequence conservation, structural genomic diversity is substantial and biologically meaningful
- High-quality, curated genomes (including the type strain) are essential for reproducible engineering and strain tracking
- ATCC resources enable reliable strain selection, comparative genomics, and industrial translation



Acknowledgements

Syracuse University College of Arts & Sciences
Forensic & National Security Sciences Institute

ATCC®

Partnerships

- CAPRA Biosciences
- Virgina Tech

Publications

- Crill JE II, et al. Complete Diploid Genome of the Type Strain *Yarrowia lipolytica* YB-423 (ATCC 18942™). J Ind Microbial Biotechnol 8: kuag002, 2026. Pub Med: 41504366
  - Published online ahead of print, January 8, 2026
  - Link is provided in the webinar console

Syracuse University

- Roy Welch
- Matt Whaley
- Zethus Avery
- Amanda Feldmann
- Code Wade
- Katherine Berge
- Michael Marciano

ATCC

- Amy Reese
- Joseph Petron
- Nikhita Puthuveetil
- Shahin Ali
- Anthony Muhle
- Rula Khairi
- Stephen King
- Kaitlyn Bentley
- Corina Tabron
- Ana Fernandes
- Briana Benton
- Jonathan Jacobs

Q & A

Syracuse University
College of Arts & Sciences
Forensic & National
Security Sciences Institute

ATCC®

Prof. James E. Crill
Professor of Practice
Syracuse University

Scott V. Nguyen, PhD
Sr. Biocuration Scientist
ATCC

Additional resources in the webinar console

- Connect with Professor Crill and Dr. Nguyen
- Access the January 8 online publication
- Search the ATCC Genome Portal
- Try searching using the NEW Taxonomy Tree in the ATCC Genome Portal
- Request a demo of ATCC Genome Portal

ATCC®

CREDIBLE LEADS TO INCREDIBLE

Thank You