

総 説

Candida albicans におけるゲノム解析

知 花 博 治 三 上 襄

千葉大学真菌医学研究センター

近年の DNA 塩基配列読み取り技術の進歩に伴い全ゲノムシーケンスが次々に報告されており、病原性真菌においても多数のゲノムシーケンスプロジェクトが進行中である。その中でスタンフォード大学において進められて来た *Candida albicans* のゲノムシーケンスプロジェクトは終了に近付いている。ここでは *C. albicans* のゲノム解析の歴史と合わせてシーケンスプロジェクトの現状について紹介する。

Key words: chromosome, physical map, whole genome shotgun, chromosome length polymorphism (CLP), assembly

病原性真菌および近縁種のゲノムシーケンスプロジェクト

現在進行中の病原性真菌およびその近縁種のゲノムシーケンスプロジェクトについて Table 1 に示した。近縁種の DNA シークエンスデータは病原性真菌の未知の遺伝子のクローニングや PCR による増幅でのプライマーのデザイン等において非常に有用なので注目したい。Table 1 には WWW 上で検索した結果と個人的な調査結果による情報が含まれている。実際にはさらに多くの菌種についてゲノムシーケンスプロジェクトが進行中と考えられる。

Candida 属に関しては *C. tropicalis* がフランス CNRS とパスツール研究所、*C. glabrata* がパスツール研究所において進行中であり、イギリスのサンガーセンターにおいては、実験的にコスミドクローンをを用いて *C. albicans* 1161 株のゲノムシーケンスを行い、約 300 kbp のシーケンスデータが GenBank に登録されている。後に詳しく述べるがスタンフォード大学においては *C. albicans* SC5314 株を用いてゲノムの 10 倍、約 160 Mbp のシーケンスが読まれた。これはゲノムのほぼ全域をカバーしている。

国別に見ると病原性真菌およびその近縁種のゲノムシーケンスプロジェクトに対する貢献度はアメリカ合衆国が最も高く 10 箇所の研究施設において 20 種 26 株、次にイギリスとフランスがそれぞれ 3 種 3 株で、日本においても長崎大学の *Aspergillus fumigatus* と *Cryptococcus neoformans*, NITE 独立行政法人製品評価技術基盤機構の *A. oryzae* の 3 種 3 株となっている。また香港大学が北京大学との共同研究により *Penicillium marneffei* のゲノムシーケンスプロジェクトを進めている。

別刷請求先: 知花 博治

〒260-8673 千葉市中央区亥鼻 1-8-1
千葉大学真菌医学研究センター

Candida albicans の倍数生

紫外線は生物の遺伝子に突然変異を生じさせ死に至らしめるが、細胞あたりのゲノム数が増加すれば対立遺伝子の数が増えるので、紫外線に対する感受性が低くなる。この現象を応用し、Olaiya and Sogin (1979) は *C. albicans* の紫外線感受性テストを行い、*Saccharomyces cerevisiae* の 1 倍体および 2 倍体との紫外線感受性を比較した。この結果 *C. albicans* の紫外線感受性は *S. cerevisiae* の 1 倍体よりはむしろ 2 倍体に近いことが明らかとなり、*C. albicans* は 2 倍体であることが推測された¹⁾。さらに Riggsby *et al.* (1982) はゲノムの DNA-DNA ハイブリダイゼーションの効率により 1 倍体あたりのゲノム DNA 量が 15-20 pg であることを導きだし、*C. albicans* の細胞あたりの DNA 量が 37 pg であることから、*C. albicans* が 2 倍体であることを支持した²⁾。

Candida albicans の電気泳動核型とゲノムサイズ

パルスフィールドゲル電気泳動法による核型の解析が 80 年代後半から 90 年代にかけて盛んに行われた³⁻¹¹⁾。*C. albicans* の核型解析がこれ程注目を浴びたのは、*C. albicans* の臨床分離株において染色体の数やサイズが異なる現象すなわち核型の多型現象が観察されたためである。ところがそのために様々な染色体の命名法が研究室ごとに存在し、混乱をきたしていた。1993 年 Chu *et al.* は制限酵素 *Sfi* I を用いて正確な物理地図を完成させ、ゲノムあたりの染色体の数は基本的に 8 本でゲノムサイズは 16 Mbp、2 倍体であることを明らかにした¹²⁾。1996 年アメリカ合衆国、サンディエゴにおける第 4 回 ASM conference *Candida* & *Candidiasis* において Chu *et al.* (1993) が提唱した染色体の番号 R, 1, 2, 3, 4, 5, 6, 7 に統一することが P.T. Magee 等によって宣言された (Fig. 1)。Chu *et al.* (1993) は *C. albicans* における多型現象が Iwaguchi *et al.* (1992) がクローニングした反復

Table 1. Genome projects for pathogenic fungi and their close relations

Beijing Genomics (China) <i>Penicillium marneffeii</i>	Univ of Aberdeen (UK) <i>Candida albicans</i> 1161
Cereon Genomics (USA) <i>Aspergillus nidulans</i>	Univ of British Columbia (Canada) <i>Cryptococcus neoformans</i> JEC 21
CNRS (France) <i>Candida tropicalis</i> CBS94	Univ of California at San Francisco (USA) <i>Histoplasma capsulatum</i> G217B
Columbia (USA) <i>Cryptococcus neoformans</i> JEC 21	Univ of Manchester (UK) <i>Aspergillus fumigatus</i> Af 293
Duke Univ (USA) <i>Cryptococcus neoformans</i> H99	Univ of Oklahoma (USA) <i>Aspergillus flavus</i>
Gene Alliance (USA) <i>Aspergillus niger</i>	<i>Aspergillus nidulans</i> FGSC A26 (biA1)
HKU-Pasteur Research Center (China) <i>Penicillium marneffeii</i>	<i>Aspergillus parasiticus</i>
Institut Pasteur (France) <i>Aspergillus fumigatus</i> Af 293	<i>Cryptococcus neoformans</i> H99
<i>Candida tropicalis</i> CBS94	<i>Cryptococcus neoformans</i> JEC 21
<i>Candida glabrata</i>	<i>Fusarium sporotrichioides</i>
Integrated Genomics Inc (USA) <i>Aspergillus niger</i> NRRL3, DSM2466	Univ of Salamanca (Spain) <i>Aspergillus fumigatus</i> Af 293
Nagasaki Univ (Japan) <i>Aspergillus fumigatus</i> Af 293	Washington Univ (USA) <i>Histoplasma capsulatum</i> G186AR
<i>Cryptococcus neoformans</i> JEC 21	<i>Histoplasma capsulatum</i> G217B
NITE (Japan) <i>Aspergillus oryzae</i> RIB40	Whitehead Inst (USA) <i>Cryptococcus neoformans</i> , serotype A
Sanger Institute (UK) <i>Aspergillus fumigatus</i> Af 293	<i>Coccidioides immitis</i>
<i>Candida albicans</i> 1161	<i>Pneumocystis carinii</i>
Stanford Univ (USA) <i>Candida albicans</i> SC5314	<i>Rhizopus arrhizus</i>
<i>Cryptococcus neoformans</i> JEC 21	<i>Aspergillus nidulans</i>
TIGR (USA) <i>Aspergillus fumigatus</i> Af 293	<i>Coccidioides posadasii/immitis</i>
<i>Coccidioides immitis</i> C735	<i>Trichophyton rubrum</i>
<i>Cryptococcus neoformans</i> JEC 21	<i>Rhizopus oryzae</i>
	<i>Aspergillus flavus</i>
	<i>Aspergillus terreus</i>
	<i>Fusarium graminearum</i>

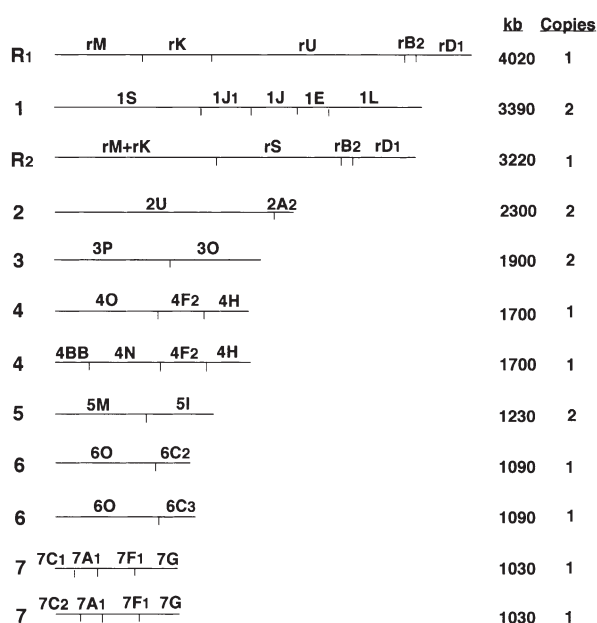


Fig. 1. The *Sfi*I map of the genome of *C. albicans* 1006. Designations of the chromosomes are shown at left, and the sizes and number of copies at right. The genome size is calculated by counting chromosomes R1 and R2 once and all the other chromosomes twice. This figure was published by Chu *et al.*¹²⁾.

配列, RPS1¹³⁾ が関係していることを示唆し, Chibana *et al.* (2000) がこれを明らかにした¹⁴⁾.

ミネソタ大学におけるゲノムプロジェクト

C. albicans のゲノムプロジェクトは1994年ミネソタ大学において Stewart Scherer および P.T. Magee によってスタートした. このとき用いられた菌株は1161株¹⁵⁾である. 最初に約4,000個の fosmid クローン¹⁶⁾ からなる *C. albicans* のゲノム DNA ライブラリーを構築し, このライブラリーを用いて Chibana *et al.* (1998) は染色体7番のテロメアを除く全域をカバーする contig map を完成させた¹⁷⁾. このとき62個のプロープ (DNA) がマッピングされ, そのうち39個についてはシークエンスが決定された. しかし contig map によって特定のプロープの順番は確定されたが染色体上での絶対距離は不明であった. そこで Random breakage mapping 法¹⁸⁾ によって絶対距離の測定が行われた. Random breakage mapping 法ではゲノム DNA に対し γ 線を照射することにより DNA を任意な位置で切断し, その後サザンハイブリダイゼーションを行うことで DNA の両端からのプロープの位置が測定できる (Fig. 2). この方法によって11個のプロープの染色体上での絶対距離が測定された. 現在他7本の染色体についても contig map を構築中であり,

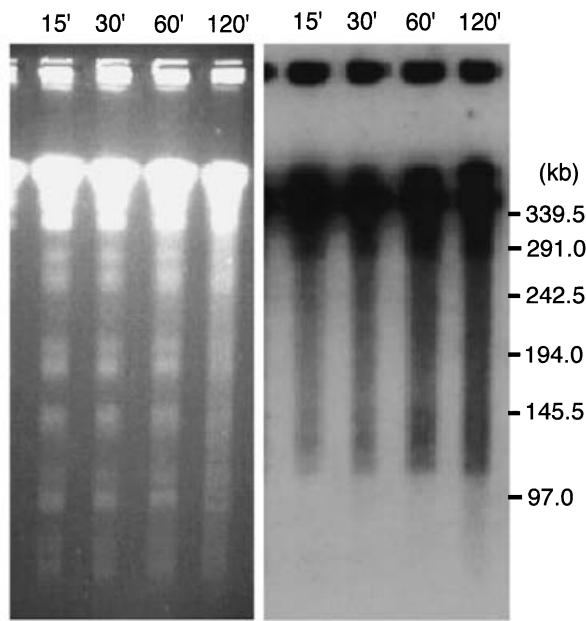


Fig. 2. Random breakage mapping. DNA was digested with *Sfi*I after irradiation in a ^{137}Cs source. The samples were loaded on a 0.9% agarose gel and run with a DRIII CHEF apparatus at 10s to 60s, 5 V/cm, 36 h, 120 degrees. A picture of the gel is shown at left side, and autoradiography of the blot hybridized with a probe from the *DFR1* gene is shown at right. Lambda ladder DNA was used as a DNA size standard. This figure was is from Chibana *et al.*¹⁷⁾.

その詳細は (<http://alces.med.umn.edu/Candida.html>) において公開されている. 物理地図の構築と平行して任意にクローニングされたゲノム DNA のシーケンスも読み取られ, その合計はゲノムの約 1.5 倍, 24 Mbp に達した. そのデータは (<http://alces.med.umn.edu/gbsearch/ybc.html>) において検索できる.

スタンフォード大学におけるゲノムプロジェクト

スタンフォード大学における *C. albicans* のゲノムシーケンスプロジェクトは 1997 年 Stewart Scherer, Ted Jones および Ron W Davis 等によってスタートした. 当時 *C. albicans* 分子生物学において用いられる菌株の主流はミネソタ大学におけるプロジェクトに用いられている 1161 株から CAI-4 株¹⁹⁾ に移っており, CAI-4 の野生株 SC5314 がスタンフォード大学におけるゲノムプロジェクトにおいて用いられた. プロジェクトではまず *C. albicans* のゲノム DNA が超音波によってランダムに切断され, 数千塩基対の DNA 断片ごとにプラスミドベクターに挿入され, 大腸菌にクローニングされた. その大腸菌からプラスミドが回収され, シークエンサーによって挿入された DNA の片側 500~700 bp の読み取りが行われた. 約 31 万個のプラスミドについてのシーケンスが読まれ, その合計は *C. albicans* のゲノムの約 10.4 倍, 156 Mbp におよんだ. その膨大なシーケンスデータは Phrap (Phil Green, <http://www.phrap.org>) を用いてシーケンスの編集 (assembling) が行われ assembly 6 にまとめられた (<http://www-sequence.stanford.edu:8080/>).

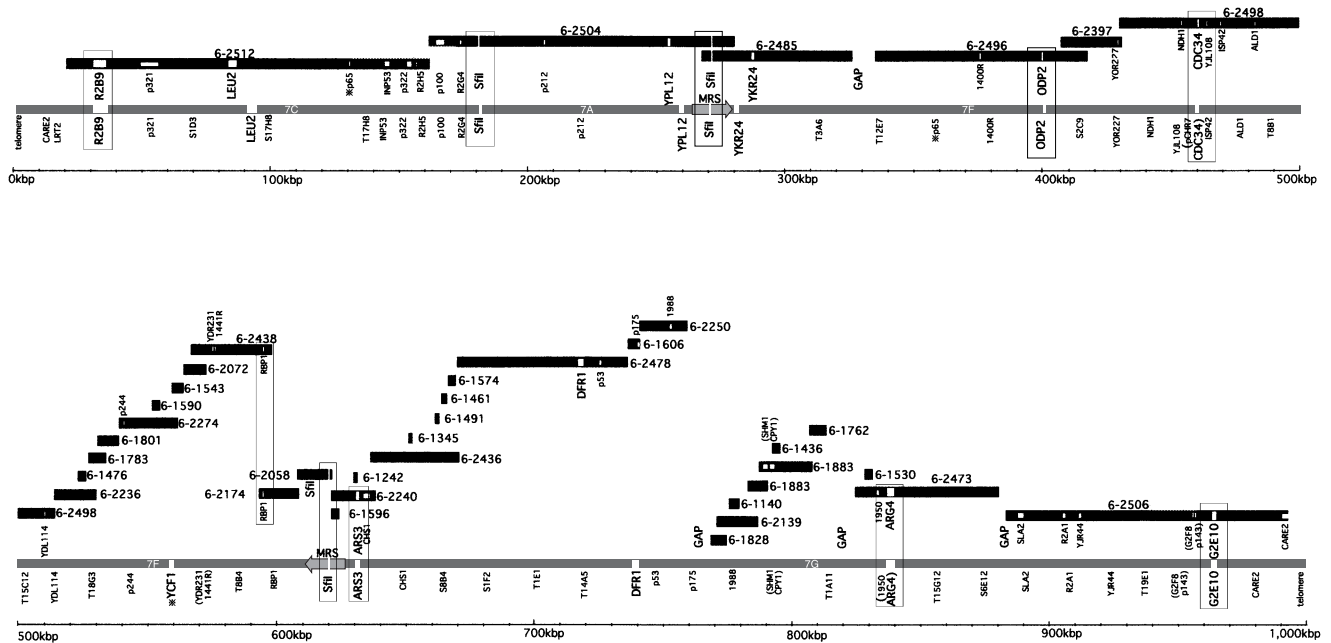


Fig. 3. Contigs from assembly 6 were mapped on a physical map for chromosome 7. The *Candida* genome sequencing project is being carried out at Stanford University, and the sequence data assembled into 1,520 contigs. The assembly is known as assembly 6 and released at <http://www-sequence.stanford>. The assembly was searched with the sequences mapped on chromosome 7, and 38 contigs were identified. Using the location of *Sfi*I sites, R2B9, RBP1, ARS1, ARG4, and G2E10 which were mapped on chromosome 7¹⁷⁾, these contigs were assigned to chromosome 7. More than 97% of chromosome 7 is covered by the contigs excluding telomeric and subtelomeric regions, and four gaps between the contigs were found.

assembly 6 のシークエンスデータは最短で 0.7 kbp, 最長で 282 kbp で 1,520 本のシークエンス (contig) で構成されているが, 染色体上の位置は不明であった。

ゲノムシークエンス完成に向けて

既に述べたようにミネソタ大学とスタンフォード大学で使われた菌株は異なっている。 *C. albicans* では核型の多型現象が観察されるが, パルスフィールドによる電気泳動核型の解析結果から両株間に染色体レベルでの違いはほとんど認められないことが分かっている (未発表)。そこで我々はゲノムシークエンスプロジェクトの進行状況を調べるために, スタンフォード大学より公開されている assembly 6 の中から染色体 7 番のシークエンスを含む contig を抽出し, ミネソタ大学において構築された染色体 7 番の物理地図上にマッピングした。シークエンスが決定されている 39 個のプロープを使って, スタンフォードのデータベース (<http://www-sequence.stanford.edu:8080/bncontigs6.html>) に対し検索し, さらに隣接する contig を walking によって検索することにより 38 本の contig が見つかった。Random breakage mapping 法によって絶対距離の測定が行われた 11 個のプロープの位置と制限酵素 *Sfi*I の位置を指標にして染色体 7 番上にマッピングを行なった (Fig. 3)。その結果, 重複していない領域 (ギャップ) が 4 箇所検出され, ギャップの合計は約 30 kbp となった。染色体の両端すなわチロメアおよびサブテロメアは複雑な反復配列を多数含んでいるので contig をこれ以上伸長するのは不可能であった。今回テロメアおよびサブテロメア領域を除くと染色体 7 番 (1 Mbp) の少なくとも 97% のシークエンスが contig でカバーされていることが明らかになった。これによりスタンフォードのゲノムシークエンスプロジェクトにおいて少なくとも 97% のゲノムシークエンスが読まれていると推測された。ゲノムの 10.4 倍の長さのシークエンスを讀んでいながら 4 つのギャップが残された理由については: (1) 大腸菌へのクローニングができない領域がある; (2) シークエンスが困難な領域がある; (3) ミネソタ大学で構築された物理地図に誤差がある; (4) スタンフォード大学で行われた assembling に間違いがある; などが考えられる。いずれにしてもギャップを埋める作業においては困難が予想されるが, 全ゲノムシークエンスの完成に向けて解析が進められることが望まれる。

Acknowledgement

We thank Stewart Scherer and Ted Jones for proofreading the contig mapping of assembly 6 on chromosome 7.

All of the *C. albicans* genome sequences of assembly 6 were provided by the Stanford Genome Technology Center prior to any publication at <http://www-sequence.stanford.edu/group/candida>. Sequencing of *C. albicans* at the Stanford Genome Technology Center was accom-

plished with the support of the NIDR and the Burroughs Wellcome Fund.

引用文献

- 1) Olaiya AF, Sogin SJ: Ploidy determination of *Candida albicans*. *J Bacteriol* **140**: 1043-1049, 1979.
- 2) Riggsby WS, Torres-Bauza LJ, Wills JW, Townes TM: DNA content, kinetic complexity, and the ploidy question in *Candida albicans*. *Mol Cell Biol* **2**: 853-862, 1982.
- 3) Snell RG, Wilkins RJ: Separation of chromosomal DNA molecules from *C. albicans* by pulsed field gel electrophoresis. *Nucleic Acids Res* **14**: 4401-4406, 1986.
- 4) Magee BB, Magee PT: Electrophoretic karyotypes and chromosome numbers in *Candida* species. *J Gen Microbiol* **133**: 425-430, 1987.
- 5) Lott TJ, Boiron P, Reiss E: An electrophoretic karyotype for *Candida albicans* reveals large chromosomes in multiples. *Mol Gen Genet* **209**: 170-174, 1987.
- 6) Lasker BA, Carle GF, Kobayashi GS, Medoff G: Comparison of the separation of *Candida albicans* chromosome-sized DNA by pulsed-field gel electrophoresis techniques. *Nucleic Acids Res* **17**: 3783-3793, 1989.
- 7) Suzuki T, Kobayashi I, Kanbe T, Tanaka K: High frequency variation of colony morphology and chromosome reorganization in the pathogenic yeast *Candida albicans*. *J Gen Microbiol* **135**: 425-434, 1989.
- 8) Iwaguchi S, Homma M, Tanaka K: Variation in the electrophoretic karyotype analysed by the assignment of DNA probes in *Candida albicans*. *J Gen Microbiol* **136**: 2433-2442, 1990.
- 9) Rustchenko-Bulgac EP: Variations of *Candida albicans* electrophoretic karyotypes. *J Bacteriol* **173**: 6586-6596, 1991.
- 10) Wickes B, Staudinger J, Magee BB, Kwon-Chung KJ, Magee PT, Scherer S: Physical and genetic mapping of *Candida albicans*: several genes previously assigned to chromosome 1 map to chromosome R, the rDNA-containing linkage group. *Infect Immun* **59**: 2480-2484, 1991.
- 11) Asakura K, Iwaguchi S, Homma M, Sukai T, Higashide K, Tanaka K: Electrophoretic karyotypes of clinically isolated yeasts of *Candida albicans* and *C. glabrata*. *J Gen Microbiol* **137**: 2531-2538, 1991.
- 12) Chu WS, Magee BB, Magee PT: Construction of an *Sfi*I macrorestriction map of the *Candida albicans* genome. *J Bacteriol* **175**: 6637-6651, 1993.
- 13) Iwaguchi S, Homma M, Chibana H, Tanaka K: Isolation and characterization of a repeated sequence (RPS1) of *Candida albicans*. *J Gen Microbiol* **138**: 1893-1900, 1992.
- 14) Chibana H, Beckerman JL, Magee PT: Fine-resolution physical mapping of genomic diversity in *Candida albicans*. *Genome Res* **10**: 1865-1877, 2000.
- 15) Goshorn AK, Grindle SM, Scherer S: Gene isolation by complementation in *Candida albicans* and applications to physical and genetic mapping. *Infect Immun* **60**: 876-884, 1992.
- 16) Kim UJ, Shizuya H, de Jong PJ, Birren B, Simon MI: Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Res* **20**: 1083-1085, 1992.
- 17) Chibana H, Magee BB, Grindle S, Ran Y, Scherer S,

- Magee PT: A physical map of chromosome 7 of *Candida albicans*. *Genetics* **149**: 1739-1752, 1998.
- 18) Game JC, Bell M, King JS, Mortimer RK: Random-breakage mapping, a rapid method for physically locating an internal sequence with respect to the ends of a DNA molecule. *Nucleic Acids Res* **18**: 4453-4461, 1990.
- 19) Fonzi WA, Irwin MY: Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* **134**: 717-728, 1993.

Genomic Analysis in *Candida albicans*

Hiroji Chibana, Yuzuru Mikami
Research Center for Pathogenic Fungi and Microbial Toxicoses, Chiba University,
1-8-1 Inohana, Chuo-ku, Chiba, 260-8673 Japan.

With the recent advances in DNA sequencing technology, a succession of entire genome sequences have been published. A number of genome projects are underway in pathogenic fungi. From these, we present the history and current status of the genomic analysis of *Candida albicans*. The sequencing project for this organism has been undertaken at Stanford University, and is now nearing the end.

この論文は、第46回日本医真菌学会総会の“シンポジウムⅢ：カンジダ属の Molecular Mycology: 臨床医にも必要な最近のトピックス”において発表されたものです。