Codon utilisation in the pathogenic yeast, *Candida albicans*

Alistair J.P. Brown, Gwyneth Bertram, Pascale J.F. Feldmann, Mark W. Pegg and Rolf K. Swoboda

Molecular and Cell Biology, University of Aberdeen, Marischal College, Aberdeen AB9 1AS, UK

Submitted May 28, 1991

The imperfect fungus, *Candida albicans* is the major yeast pathogen in humans (1). We have analysed the codon usage in eleven *C. albicans* gene sequences reported in the EMBL and GenBank databases (Table). The codons are divided into three groups (low, medium and high) based upon their codon usage (Table), and these groups are compared with analogous groups for the highly expressed genes from *S. cerevisiae* and the divergent yeast, *Schizosaccharomyces pombe* (12).

The subsets of strongly preferred codons in each organism overlap (for example, with the codons GGT, GTT, TTG and CAA). However, significant differences exist between these subsets, and for *C. albicans* these differences represent a strong preference for an A or T in the third, or 'wobble' position. For example, the preferred codon for lysine is AAA in *C. albicans*, but is AAG in *S. cerevisiae* and *S. pombe*, and the leucine codon TTA is frequently used in *C. albicans*, but not in the other two yeasts (Table). This should be considered when using amino acid sequence data for the design of synthetic oligonucleotide probes to isolate *C. albicans* genes.

The bias towards an A or T in the wobble position may reflect the high A:T content of *C. albicans* DNA (65.3% in the genes analysed here). However, the preference for the termination codon TAA is probably due to the efficiency of natural suppression of termination by TGA and TAG in *C. albicans* (13).

The CTG codon, which normally encodes leucine, has been shown to encode serine in *C. cylindracea* (14). Since to our knowledge there has been no analogous report for *C. albicans*, no reference to this is made in the Table. Such a change in tRNA specificity would prevent the use of heterologous reporter genes in *C. albicans*.

ACKNOWLEDGEMENTS

DNA sequence analysis was performed on SEQNET at the SERC Daresbury Laboratory. RKS and GB are supported by the Wellcome Trust, and MWP and PJFF are supported by SERC and Glaxo Group Research Studentships, respectively.

REFERENCES

1. Odds, F.C. (1988) *Candida and candidosis: a review and bibliography* (2nd edition) Bailliere Tindall, London.
2. Smith, H.A., Allaudeen, H.S., Whitman, M.H., Koltin, Y. and Gorman, J.A. (1988) *Gene* 63, 53–63.
3. Losberger, C. and Ernst, J.F. (1989) *Curr. Genet.* 16, 153–157.
4. Losberger, C. and Ernst, J.F. (1989) *Nucl. Acids Res.* 17, 9488.
5. Denich, K., Malloy, P.J. and Feldman, D. (1990) GenBank M54910.

6. Saporito, S.M. and Sypherd, P.S. (1991) EMBL M61128.
7. Koser, P.L., Livi, G.P., Levy, M.A., Rosenberg, M. and Bergsma, D.J. (1990) *Gene* 96, 189–195.
8. Lai, M.H. and Kirsch, D.R. (1989) *Nucl. Acids Res.* 17, 804.
9. Lott, T.J., Boiron, P., Page, L.S., Benson, J. and Reiss, E. (1989) *Nucl. Acids Res.* 17, 1779.
10. Sundstrom, P., Smith, D. and Sypherd, P.S. (1990) *J. Bacteriol.* 172, 2036–2045.
11. Singer, S.C., Richards, C.A., Ferone, R., Benedict, D. and Ray, P. (1989) *J. Bacteriol.* 171, 1372–1378.
12. Sharp, P.M., Cowe, E., Higgins, D.G., Shields, D.C., Wolfe, K.H. and Wright, F. (1988) *Nucl. Acids Res.* 16, 8207–8211.
13. Santos, M., Colthurst, D.R., Wills, N., McLaughlin, C.S. and Tuite, M.F. (1990) *Curr. Genet.* 17, 487–491.
14. Kawaguchi, Y., Honda, H., Taniguchi-Morimura, J. and Iwasaki, S. (1989) *Nature* 341, 164–166.

Table. Codon usage in *Candida albicans*.

<i>C. albicans</i>			Preferred Codons			
A.Acid	Codon	Number [a]	Fraction [b]	C.s. [c]	S.c. [d]	S.p. [d]
Gly	GGA	24	0.08	-	-	-
Gly	GGC	23	0.07	-	-	-
Gly	GGA	19	0.04	-	-	-
Gly	GGC	23	0.08	-	-	-
Gly	GGA	23	0.08	-	-	-
Arg	AGA	24	0.08	-	-	-
Arg	AGC	24	0.08	-	-	-
Val	GTA	18	0.07	-	-	-
Val	GTC	17	0.06	-	-	-
Val	GTA	59	0.22	-	-	-
Ala	GCA	2	0.01	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	-
Ala	GCC	18	0.06	-	-	-
Ala	GCA	18	0.06	-	-	