destabilization of the H1 helix and, for ΔA296/A301U, stabilization of the SD hairpin (*29*), both of which should favor the monomer. Mutations in AUG that inhibit genome dimerization and suppress packaging (*30*, *31*) are expected to disrupt base pairing in the U5:AUG helix and ψ-hairpin stem, thereby destabilizing the tandem three-way junction structure required for Gag binding. In vitro splicing activity is also attenuated by dimerization (*12*, *32*), and this can be attributed to sequestration of the major splice-site recognition sequence within the three-way-2 junction. Antisense oligonucleotides with complementarity to the SD loop inhibit dimerization, and this is likely due to their ability to competitively block formation of the tandem three-way junction (*25*).

The Ψ^{CES} structure also explains the exquisite selectivity of HIV-1 to package its unspliced genome (*1*, *2*). Residues immediately downstream of the major splice site are base-paired within the H1 helix and are thus integral to the formation of the tandem three-way junction structure. Although unspliced and spliced HIV-1 mRNAs contain identical 5′ sequences (G1 to G289), differences in spliced mRNA sequences derived from 3′ exons would preclude formation of the packaging competent junction structure. Similarly, because SD appears to exist as a hairpin in the monomeric, unspliced 5′ leader (*12*), it is likely that monomeric genomes are also ignored during virus assembly because they do not adopt the tandem three-way junction structure.

Compared with the proteins of HIV-1, structural information for the viral nucleic acids is sparse. RNAs in general are vastly underrepresented in the structural data banks (99,000 proteins versus 2700 RNA structures), partly because of NMR technical challenges and difficulties obtaining suitable crystals for x-ray diffraction (*19*, *20*). The fr-RNA ^{2}H-edited NMR approach enables efficient segmental labeling without requiring enzymatic ligation. Given the ubiquity of hairpin elements that can serve as fragmentation or annealing sites, this method should be generally applicable to modest-sized RNAs (~160 nucleotides).

REFERENCES AND NOTES

1. J. M. Coffin, S. H. Hughes, H. E. Varmus, Eds., *Retroviruses* (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1997).
2. R. Berkowitz, J. Fisher, S. P. Goff, *Curr. Top. Microbiol. Immunol.* **214**, 177–218 (1996).
3. J. Chen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13535–13540 (2009).
4. K. Lu, X. Heng, M. F. Summers, *J. Mol. Biol.* **410**, 609–633 (2011).
5. N. Jouvenet, S. M. Simon, P. D. Bieniasz, *J. Mol. Biol.* **410**, 501–511 (2011).
6. M. Kuzembayeva, K. Dilley, L. Sardo, W.-S. Hu, *Virology* **454-455**, 362–370 (2014).
7. S. Schwartz, B. K. Felber, D. M. Benko, E. M. Fenyö, G. N. Pavlakis, *J. Virol.* **64**, 2519–2529 (1990).
8. O. A. Nikolaitchik *et al.*, *PLOS Pathog.* **9**, e1003249 (2013).
9. A. M. Lever, *Adv. Pharmacol.* **55**, 1–32 (2007).
10. J.-C. Paillart, M. Shehu-Xhilaga, R. Marquet, J. Mak, *Nat. Rev. Microbiol.* **2**, 461–472 (2004).
11. J. Greatorex, *Retrovirology* **1**, 22 (2004).
12. T. E. M. Abbink, B. Berkhout, *J. Virol.* **82**, 3090–3098 (2008).
13. K. Lu *et al.*, *Science* **334**, 242–245 (2011).
14. T. E. M. Abbink, B. Berkhout, *J. Biol. Chem.* **278**, 11601–11611 (2003).
15. X. Heng *et al.*, *J. Mol. Biol.* **417**, 224–239 (2012).
16. K. A. Wilkinson *et al.*, *PLOS Biol.* **6**, e96 (2008).
17. S. B. Kutluay *et al.*, *Cell* **159**, 1096–1109 (2014).
18. Information on materials and methods is available on *Science* Online.
19. K. Lu, Y. Miyazaki, M. F. Summers, *J. Biomol. NMR* **46**, 113–125 (2010).
20. O. Duss, P. J. Lukavsky, F. H. Allain, *Adv. Exp. Med. Biol.* **992**, 121–144 (2012).
21. L. M. Koharudin, A. M. Bonvin, R. Kaptein, R. Boelens, *J. Magn. Reson.* **163**, 228–235 (2003).
22. S. Barton, X. Heng, B. A. Johnson, M. F. Summers, *J. Biomol. NMR* **55**, 33–46 (2013).
23. F. J. van Hemert, A. C. van der Kuyl, B. Berkhout, *RNA Biol.* **10**, 211–215 (2013).
24. J. C. Kenyon, L. J. Prestwood, S. F. Le Grice, A. M. Lever, *Nucleic Acids Res.* **41**, e174 (2013).
25. J. Deforges, N. Chamond, B. Sargueil, *Biochimie* **94**, 1481–1489 (2012).
26. T. L. South, M. F. Summers, *Protein Sci.* **2**, 3–19 (1993).
27. R. N. De Guzman *et al.*, *Science* **279**, 384–388 (1998).
28. F. Baudin *et al.*, *J. Mol. Biol.* **229**, 382–397 (1993).
29. T. E. Abbink, M. Ooms, P. C. Haasnoot, B. Berkhout, *Biochemistry* **44**, 9058–9066 (2005).
30. D. T. Poon, E. N. Chertova, D. E. Ott, *Virology* **293**, 368–378 (2002).
31. O. Nikolaitchik, T. D. Rhodes, D. Ott, W.-S. Hu, *J. Virol.* **80**, 4691–4697 (2006).
32. J. A. Jablonski, E. Buratti, C. Stuani, M. Caputi, *J. Virol.* **82**, 8038–8050 (2008).

ACKNOWLEDGMENTS

This research was supported by grants from the National Institute of General Medical Sciences (NIGMS, R01 GM42561 to M.F.S. and A.T. and P50 GM 103297 to M.S., B.J., and D.A.C.). S.B., N.C.B., and S.M. were supported by a NIGMS grant for enhancing minority access to research careers (MARC U*STAR 2T34 GM008663), and S.B., J.S., N.C.B., and S.M. were supported by an HHMI undergraduate education grant. We thank the HHMI staff at UMBC for technical assistance and B. Rife (University of Florida) for advice regarding the phylogenetic analysis. The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, NIH: pNL4-3 from M. Martin. Atomic coordinates have been deposited into the Protein Data Bank with accession code 2N1Q. NMR chemical shifts and restraints have been deposited into the Biological Magnetic Resonance Bank with accession code 25571.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6237/917/suppl/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 and S2
References (*33*–*52*)

14 February 2015; accepted 22 April 2015
10.1126/science.aaa9266

EVOLUTION

Systematic humanization of yeast genes reveals conserved functions and genetic modularity

Aashiq H. Kachroo,[1] **Jon M. Laurent,**[1] **Christopher M. Yellman,**[1] **Austin G. Meyer,**[1,2] **Claus O. Wilke,**[1,2,3] **Edward M. Marcotte**[1,2,4]*

To determine whether genes retain ancestral functions over a billion years of evolution and to identify principles of deep evolutionary divergence, we replaced 414 essential yeast genes with their human orthologs, assaying for complementation of lethal growth defects upon loss of the yeast genes. Nearly half (47%) of the yeast genes could be successfully humanized. Sequence similarity and expression only partly predicted replaceability. Instead, replaceability depended strongly on gene modules: Genes in the same process tended to be similarly replaceable (e.g., sterol biosynthesis) or not (e.g., DNA replication initiation). Simulations confirmed that selection for specific function can maintain replaceability despite extensive sequence divergence. Critical ancestral functions of many essential genes are thus retained in a pathway-specific manner, resilient to drift in sequences, splicing, and protein interfaces.

The ortholog-function conjecture posits that orthologous genes in diverged species perform similar or identical functions (*1*). The conjecture is supported by comparative analyses of gene-expression patterns, genetic interaction maps, and chemogenomic profiling (*2*–*6*), and it is widely used to predict gene function across species. However, even if two genes perform similar functions in different organisms, it may not be possible to replace one for the other, particularly if the organisms are widely diverged. The extent to which deeply divergent orthologs can stand in for each other, and which principles govern such functional equivalence across species, is largely unknown.

In this study, we systematically addressed these questions by replacing a large number of yeast genes with their human orthologs. Humans and the baker's yeast *Saccharomyces cerevisiae* diverged from a common ancestor approximately 1 billion years ago (*7*). They share several thousand

---

[1]Center for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA. [2]Center for Computational Biology and Bioinformatics, University of Texas at Austin, Austin, TX 78712, USA. [3]Department of Integrative Biology, University of Texas at Austin, Austin, TX 78712, USA. [4]Department of Molecular Biosciences, University of Texas at Austin, Austin, TX 78712, USA.
*Corresponding author. E-mail: marcotte@icmb.utexas.edu

Downloaded from www.sciencemag.org on May 21, 2015

orthologous genes, accounting for more than one-third of the yeast genome (*8*). Yeast and human orthologs tend to be recognizable but often highly diverged; amino acid identity ranges from 9 to 92%, with a genome-wide average of 32%. Although we know of individual examples of human genes capable of replacing their fungal orthologs (*9*–*12*), the extent and specific conditions under which human genes can substitute for their yeast orthologs are generally not known.

We focused on the set of genes essential for yeast cell growth under standard laboratory conditions (*13*, *14*) and for which the yeast-human orthology is 1:1 (i.e., genes without lineage-specific duplicate genes that might mask the effects). Based on the availability of full-length human cDNA recombinant clones (*15*, *16*) and matched yeast strains with conditionally null alleles of the test genes (*17*–*19*), we selected 469 human genes to study (Fig. 1A).

We first subcloned and sequence-verified each human protein coding sequence into a single-copy, centromeric yeast plasmid under the transcriptional control of either an inducible (GAL) or constitutively active (GPD) promoter (see supplementary materials and methods). We assembled a matched set of yeast strains in which each orthologous yeast gene could be conditionally down-regulated [via a tetracycline-repressible promoter (*17*)], inactivated [via a temperature-sensitive allele (*18*)], or segregated away genetically [following sporulation of a heterozygous diploid deletion strain (*13*, *19*)] (Fig. 1A and fig. S1). After verifying that the loss of the relevant yeast gene conferred a strong growth defect, we tested whether expression of the human ortholog could complement the growth defect, as illustrated for several examples in Fig. 1B (also figs. S2 to S4). When expressed in the permissive condition, 73 of the human genes exhibited toxicity; reducing the genes' expression levels allowed us to assay replacement in 66 cases (table S1).

Overall, we performed 652 informative growth assays surveying 414 human-yeast orthologs (Fig. 1, A and C). In total, 176 yeast genes (43%) could be replaced by their human orthologs in at least one of the three strain backgrounds, whereas 238 (57%) could not (table S1). We collated previously published reports of yeast gene complementation by human genes: Our assays recapitulated these cases with 90% precision and 72% recall (table S1), and incorporating the literature data for subsequent analyses brought the observed complementation rate to 47% (Fig. 1C). For randomly selected subsets of strains, we additionally validated the assays by confirming human protein expression using Western blot analysis (fig. S5), verifying complementation by tetrad dissection (table S1), and subcloning the yeast test genes into the assay vectors and confirming positive complementation (table S2).

Given that roughly half of the tested human genes successfully replaced and half did not, we next investigated factors determining replaceability. We assembled 104 quantitative features of the genes or ortholog pairs, including calculated properties of the genes' sequences (e.g., gene and protein lengths, sequence similarities, codon usage, and predicted protein aggregation potential) and properties such as protein interactions, mRNA and protein abundances, transcription and translation rates, and mRNA splicing features (table S3). We then quantified how well each feature predicted replaceability (Fig. 2A and table S3).

Notably, sequence similarity only partly predicted replaceability. This tendency was strongest for highly similar (>50% amino acid identity) or dissimilar (<20%) ortholog pairs. However, most pairs fell into an intermediate range of 20 to 50% sequence identity, which poorly predicted replaceability (Fig. 2B). Instead, replaceability was best predicted by properties of specific gene modules. In particular, proteins in the same pathway or complex tended to be similarly replaceable (Fig. 2A). Replaceable genes also tended to be shorter and more highly expressed. Using these features in a supervised Bayesian network classification algorithm (fig. S6), we achieved a high overall cross-validated prediction rate [area under the receiver operating characteristic (ROC) curve of 0.825 (Fig. 2A)] and correct prediction of 8 of 10 literature cases withheld from all computational analyses (table S4). Properties such as human gene splice-form counts, yeast 5′ and 3′ untranslated region lengths, codon adaptation measures, and yeast mRNA

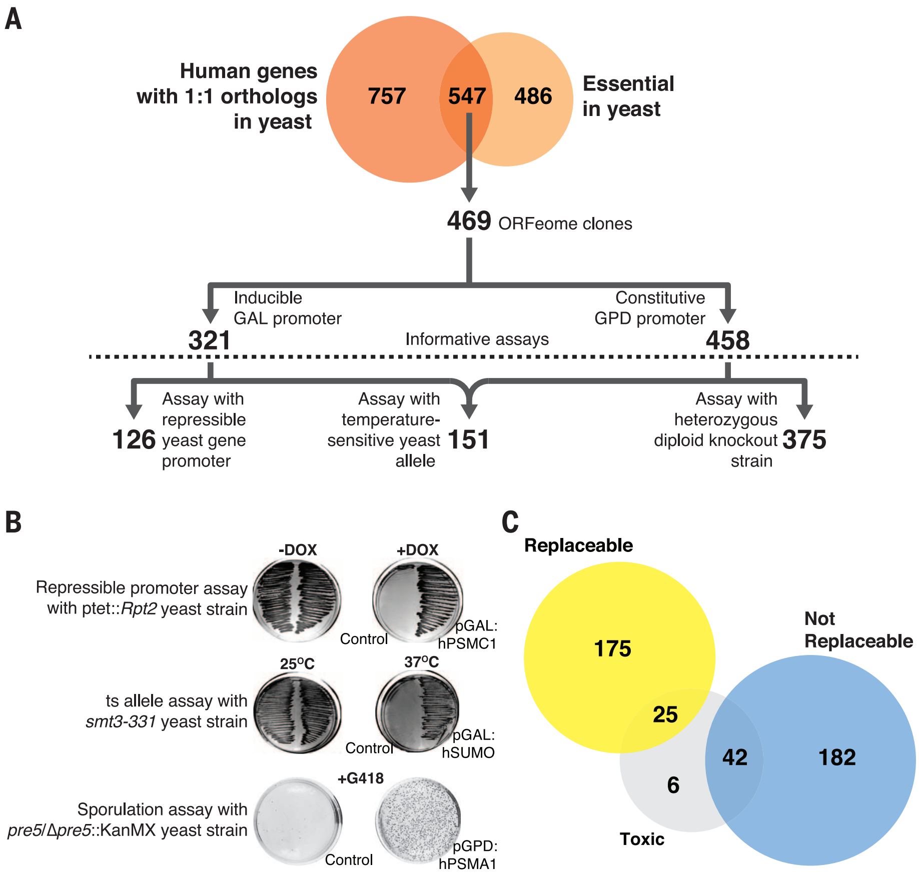

A
Human genes with 1:1 orthologs in yeast
757
547
486
Essential in yeast
469 ORFeome clones
Inducible GAL promoter
Constitutive GPD promoter
321
Informative assays
458
Assay with repressible yeast gene promoter
126
Assay with temperature-sensitive yeast allele
151
Assay with heterozygous diploid knockout strain
375
B
-DOX
+DOX
Repressible promoter assay with ptet::*Rpt2* yeast strain
Control
pGAL: hPSMC1
25°C
37°C
ts allele assay with *smt3-331* yeast strain
Control
pGAL: hSUMO
+G418
Sporulation assay with *pre5*/Δ*pre5*::KanMX yeast strain
Control
pGPD: hPSMA1
C
Replaceable
175
25
6
42
182
Not Replaceable
Toxic

Fig. 1. Systematic functional replacement of essential yeast genes by their human counterparts. (**A**) Of 547 human genes with 1:1 orthology to essential yeast genes, 469 human open reading frames (ORFs) were subcloned into single-copy yeast expression vectors under control of either the GAL or GPD promoters. Using three distinct assay classes (repressible yeast-gene promoter, temperature-sensitive yeast allele, and heterozygous diploid knockout strain), we obtained 126, 151, and 375 informative replaceability assays, respectively. (**B**) Representative examples of the three assay classes. (**C**) Combining assays and literature, 200 human genes could functionally replace their yeast orthologs and 224 genes could not. Some human genes were toxic using GAL induction but replaced their yeast orthologs upon reducing expression.

half-lives showed little relationship with replaceability (Fig. 2A and table S3).

The strong association between replaceability and gene modules led us to investigate this phenomenon in more depth, examining replaceability as a function of specific protein complexes and pathways. Broad Kyoto Encyclopedia of Genes and Genomes (KEGG) (*20*) pathway classes showed highly differential replaceability: Metabolic enzymes (e.g., enzymes participating in lipid, amino acid, and carbohydrate metabolism) tended to be replaceable, whereas proteins involved in DNA replication and repair or in cell growth tended not to be replaceable (Fig. 2C).

Among large protein complexes and pathways, we observed both extremes of replaceability. Some were entirely nonreplaceable: For example, we did not observe a single successful replacement among 13 tested members of the TriC chaperone complex, the DNA replication initiation origin recognition complex, or its interacting minichromosome maintenance (MCM) complex (Fig. 3, A and B). In contrast, some pathways were almost entirely replaceable: Among 19 components of the sterol biosynthesis pathway (which catalyzes the conversion of acetyl–coenzyme A to cholesterol in humans and ergosterol in yeast), only the human farnesyl-diphosphate farnesyltransferase 1 (*FDFT1*) and farnesyl diphosphate synthase (*FDPS*) enzymes failed to replace their yeast orthologs. All other tested components were replaceable, suggesting that yeast and humans both retain the same essential complement of ancestral sterol biosynthesis functionality (Fig. 3C and fig. S7).

The modular nature of replaceability was particularly evident in the case of the 26*S* proteasome complex. Of 28 tested subunits, 21 human

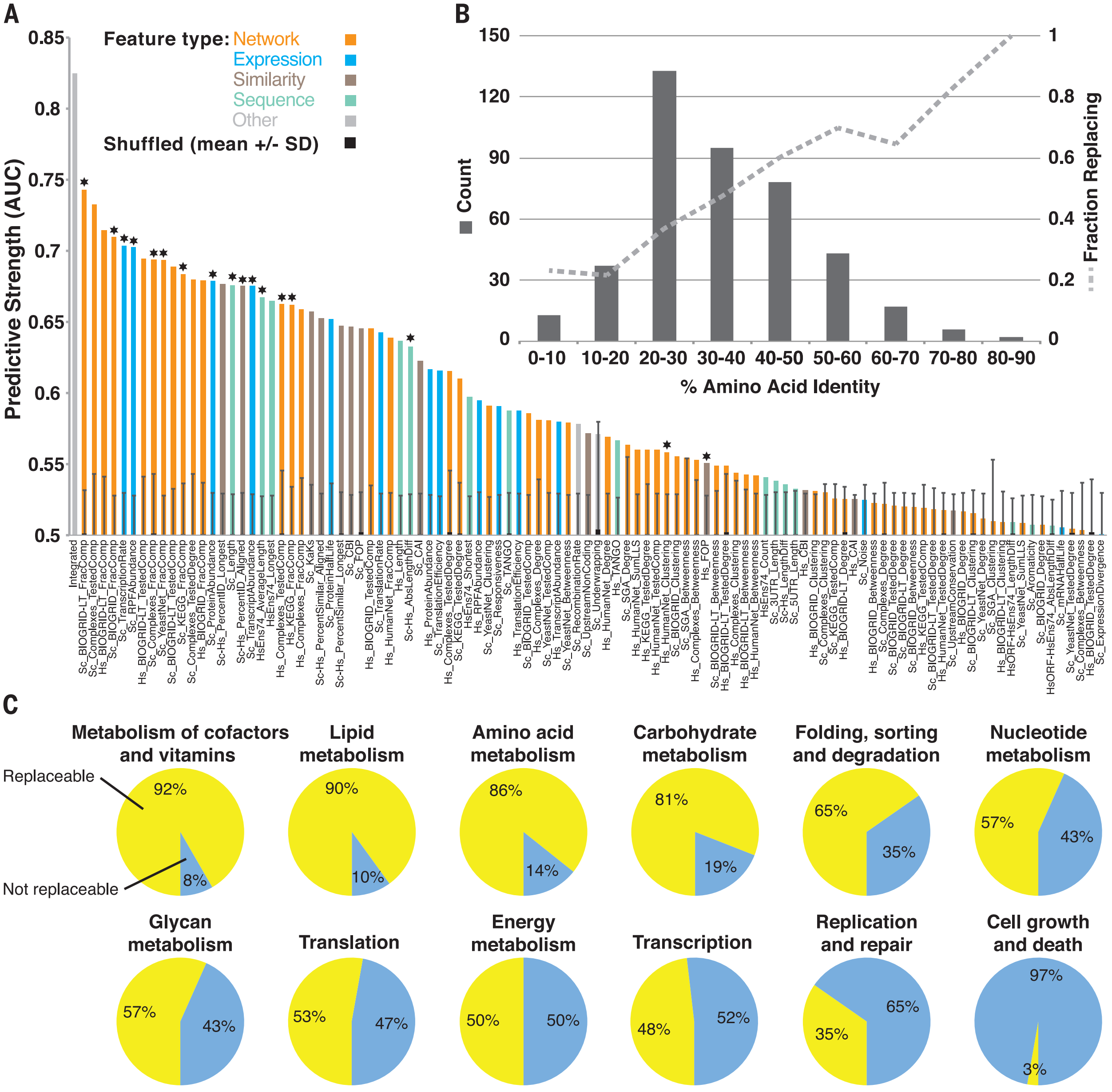
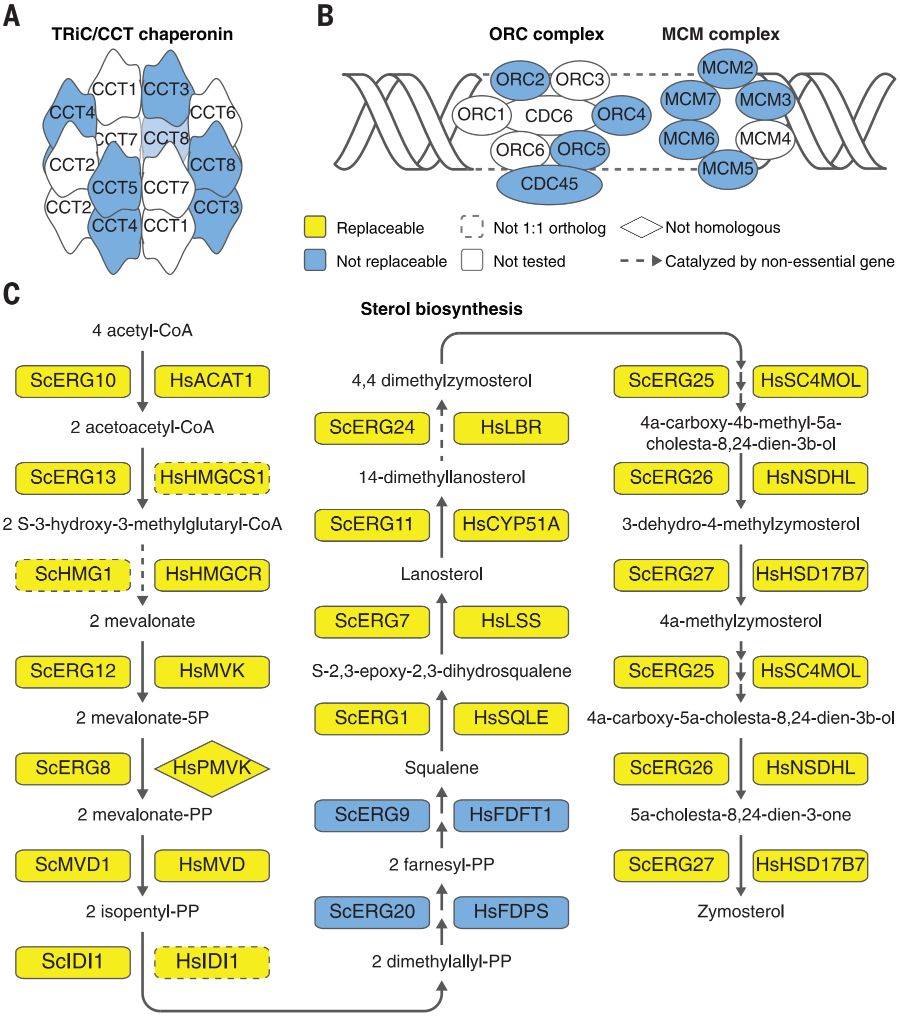


Fig. 2. Properties of gene modules can predict replaceability. (**A**) One hundred four quantitative features of proteins or ortholog pairs were evaluated for their ability to explain replaceability, assessing each feature's predictive strength as the area under a ROC curve (AUC) and determining significance by shuffling replacement status 1000 times, measuring mean AUCs ± 1 SD (error bars). AUCs above 0.58 were generally individually significant with 95% confidence. Starred features were included in the integrated classifier (leftmost bar). (**B**) Distribution of amino acid identities among the tested ortholog pairs (left *y* axis) and fraction of replaceable genes in each sequence-identity bin (right *y* axis). (**C**) Relative proportion of replaceable and nonreplaceable genes among 12 broad KEGG (*20*) pathway classes.

Fig. 3. The modular nature of functional replacement. (**A**) None of the four tested human TRiC/CCT chaperonin genes replaced their yeast counterparts. (**B**) Similarly, no genes tested in the origin recognition complex (ORC) or the MCM complex were replaceable. (**C**) In contrast, 17 of 19 sterol biosynthesis genes were replaceable. In two cases, the yeast gene had two human orthologs but only one could complement. Human *HMGCS1* (but not *HMGCS2*) replaced yeast *ERG13*; human *IDI1* (but not *IDI2*) replaced yeast *IDI1*. Human PMVK, a nonhomologous protein that carries out the same reaction as yeast Erg8 (*27*), complemented temperature-sensitive allele *erg8-1*.


A
TRiC/CCT chaperonin
CCT1
CCT3
CCT4
CCT6
CCT7
CCT8
CCT2
CCT8
CCT5
CCT7
CCT2
CCT3
CCT4
CCT1
B
ORC complex
MCM complex
ORC2
ORC3
ORC1
CDC6
ORC4
ORC6
ORC5
CDC45
MCM2
MCM7
MCM3
MCM6
MCM4
MCM5
Replaceable
Not replaceable
Not 1:1 ortholog
Not tested
Not homologous
Catalyzed by non-essential gene
C
Sterol biosynthesis
4 acetyl-CoA
ScERG10
HsACAT1
2 acetoacetyl-CoA
ScERG13
HsHMGCS1
2 S-3-hydroxy-3-methylglutaryl-CoA
ScHMG1
HsHMGCR
2 mevalonate
ScERG12
HsMVK
2 mevalonate-5P
ScERG8
HsPMVK
2 mevalonate-PP
ScMVD1
HsMVD
2 isopentyl-PP
ScIDI1
HsIDI1
2 dimethylallyl-PP
ScERG20
HsFDPS
2 farnesyl-PP
ScERG9
HsFDFT1
Squalene
ScERG1
HsSQLE
S-2,3-epoxy-2,3-dihydrosqualene
ScERG7
HsLSS
Lanosterol
ScERG11
HsCYP51A
14-dimethyllanosterol
ScERG24
HsLBR
4,4 dimethylzymosterol
ScERG25
HsSC4MOL
4a-carboxy-4b-methyl-5a-cholesta-8,24-dien-3b-ol
ScERG26
HsNSDHL
3-dehydro-4-methylzymosterol
ScERG27
HsHSD17B7
4a-methylzymosterol
ScERG25
HsSC4MOL
4a-carboxy-5a-cholesta-8,24-dien-3b-ol
ScERG26
HsNSDHL
5a-cholesta-8,24-dien-3-one
ScERG27
HsHSD17B7
Zymosterol

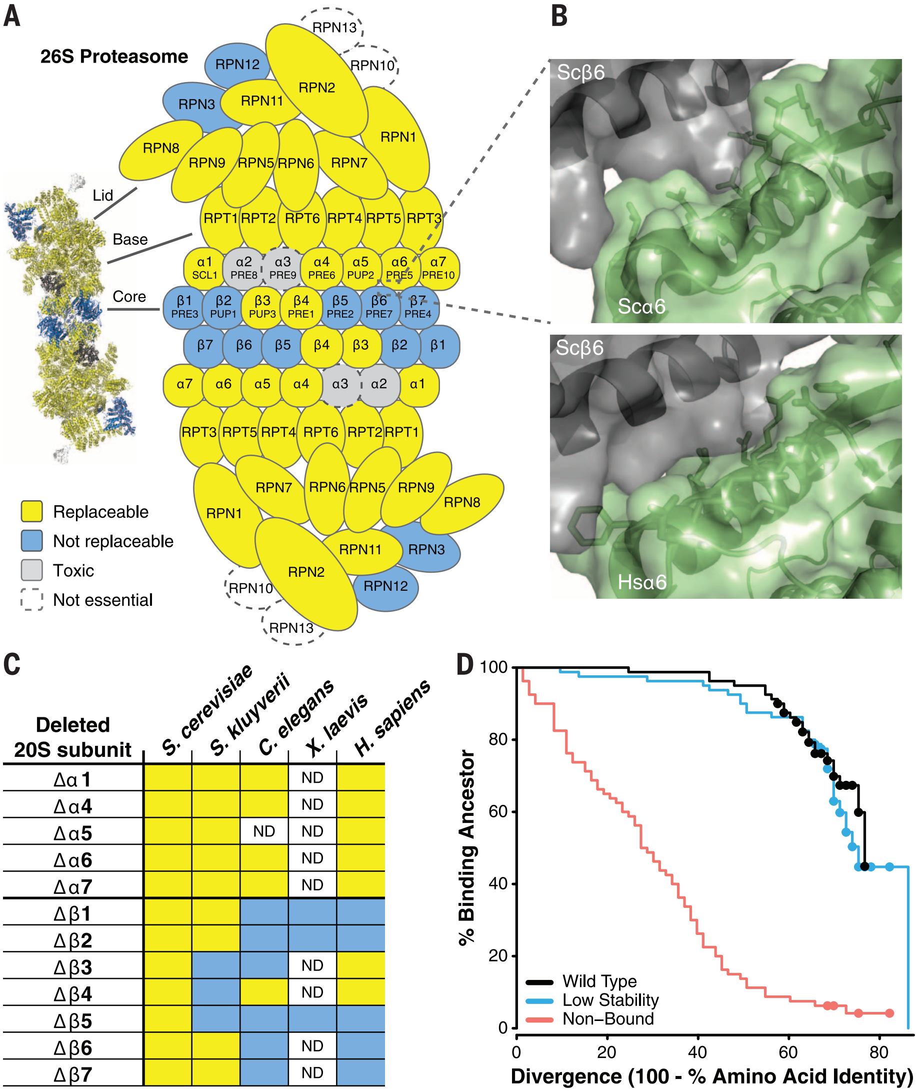
genes replaced their yeast counterparts (Fig. 4A). However, the nonreplaceable subunits were not randomly distributed; rather, they clustered in two physically interacting groups: one consisting of the 19*S* lid components Rpn3 and Rpn12 and one consisting of the 20*S* inner-core heptameric beta-ring subunits β1, β2, β5, β6, and β7. Thus, of the two central heteroheptameric rings, all testable components of the alpha ring replaced, whereas most of the beta ring did not.

An examination of the alpha and beta subunit structures showed that subunit-subunit interfacial amino acids were conserved to similar degrees between yeast and human subunits (fig. S8A), although beta subunits exhibited elevated rates of nonsynonymous substitutions compared with alpha subunits (fig. S8B). Even when interfacial amino acids were only partly conserved, modeling human alpha subunits into the known structure of the yeast proteasome (*21*) revealed that human proteins could be sterically accommodated into the yeast intersubunit interface, as shown for human α6 (Fig. 4B) packing against yeast β6, in spite of sharing only 50% identical amino acids at the interface (fig. S8A). Only orthologous alpha subunits replaced; nonorthologs failed (fig. S9).

We further confirmed this trend across alpha and beta proteasome subunits by cloning and assaying subunits from additional organisms, including another yeast (*Saccharomyces kluveri*), the nematode *Caenorhabditis elegans*, and several beta subunits from the frog *Xenopus laevis*. In all cases, alpha subunits complemented the loss of the yeast orthologs, whereas beta subunits generally failed to complement (Fig. 4C). The pattern of replaceability across species suggests that alpha and beta subunits experienced different evolutionary pressures, in each case operating at the level of the system of genes (the alpha or beta heteroheptamer).

To determine further why proteasome alpha subunits were replaceable but beta subunits were not, we isolated human β2 subunit mutants that complemented the yeast defect (figs. S10 to S12). A single serine-to-glycine substitution [Ser^{214}→Gly^{214} (S214G)] was sufficient to rescue growth (fig. S11). β2 subunits act as proteases, but yeast β2 catalytic activity is dispensable if the proteasome assembles with other functioning protease subunits (*22*). Notably, a catalytically dead [Thr^{44}→ Ala^{44} (T44A)] human β2 failed to complement, whereas an S214G, T44A double mutant complemented successfully (fig. S11). We conclude that the S214G mutant is competent to assemble an intact proteasome, although the subunit may not be catalytically active. Thus, native human β2 needs only one amino acid change to pack within the yeast proteasome.

Theory predicts that evolutionary divergence creates Dobzhansky-Muller incompatibilities, because evolutionarily novel mutations in one species are untested in the other species' genetic background and may be deleterious there (*23*, *24*). To better understand how proteins retain the ability to interact with their ortholog's interaction partners, even when they have diverged substantially, we developed a biochemically realistic divergence model in which we simulated the evolution of two physically interacting proteins, which both diverge over time. We considered three distinct scenarios: (i) Both thermodynamic stability and binding to the extant partner were selected at ancestral levels; (ii) binding was selected at ancestral levels but stability was not; and (iii) stability was selected at ancestral levels but binding was not. Thermodynamic stability ($\Delta G^{\text{folding}}$) and binding energy ($\Delta G^{\text{interaction}}$) were calculated using the empirical FoldX energy function (*25*). Under all scenarios, we evaluated whether an evolved member of the pair could still bind to its ancestral partner,



Deleted 20S subunit	*S. cerevisiae*	*S. kluyverii*	*C. elegans*	*X. laevis*	*H. sapiens*
Δα**1**	🟨	🟨	🟨	ND	🟨
Δα**4**	🟨	🟨	🟨	ND	🟨
Δα**5**	🟨	🟨	ND	ND	🟨
Δα**6**	🟨	🟨	🟨	ND	🟨
Δα**7**	🟨	🟨	🟨	ND	🟨
Δβ**1**	🟨	🟨	🟦	🟦	🟦
Δβ**2**	🟨	🟨	🟦	🟦	🟦
Δβ**3**	🟨	🟦	🟦	ND	🟨
Δβ**4**	🟨	🟦	🟨	ND	🟨
Δβ**5**	🟨	🟦	🟦	🟦	🟦
Δβ**6**	🟨	🟨	🟦	ND	🟦
Δβ**7**	🟨	🟨	🟦	ND	🟦

Fig. 4. Proteasome subunits are differentially replaceable. (**A**) Yeast 26*S* proteasome genes were generally replaceable, except for two interacting clusters, in the 19*S* regulatory "lid" particle and in the 20*S* core β-subunit ring. (**B**) The yeast α6-β6 subunit interface (top panel) sterically accommodates the human subunit (bottom panel, showing superposition of human α6 onto the yeast α6) despite 50% sequence identity at the interface. (**C**) Alpha subunits from diverse eukaryotes generally complemented the yeast mutant, but beta subunits did not (unlike plasmid-expressed *S. cerevisiae* genes, included as positive controls). ND, not determined. (**D**) In simulated evolution of interacting proteins Ubc9 and Smt3, if binding to the extant partner is not enforced ("Non-Bound"), a protein's ability to bind its ancestral partner decays rapidly as sequences diverge. However, if extant binding is enforced ("Wild Type" and "Low Stability"), even highly diverged proteins often still bind to their ancestral partners. (Dots indicate right-censored data; see fig. S14.)

for which binding was not enforced. We found that ancestral binding decayed rapidly under scenario (iii) but much more slowly under the other two scenarios (Fig. 4D and figs. S13 to S15). Natural selection for a protein interaction thus preserves the interaction interface in a manner consistent with binding to the ancestral partner (figs. S16 and S17), even though many lineages will eventually accumulate mutations that cause incompatibilities with the ancestral interactor.

Our data demonstrate that a substantial portion of conserved yeast and human genes perform much the same roles in both organisms, to an extent that the protein-coding DNA of a human gene can actually substitute for that of the yeast. The strong pathway-specific pattern of individual replacements suggests that group-wise replacement of the genes should be feasible, raising the possibility of humanizing entire cellular processes in yeast. Such strains would simplify drug discovery against human proteins, enable studies of the consequences of human genetic polymorphisms [as in (*26*) and fig. S7], and empower functional studies of entire human cellular processes in a simplified organism.

REFERENCES AND NOTES

1. T. Gabaldón, E. V. Koonin, *Nat. Rev. Genet.* **14**, 360–366 (2013).
2. N. L. Nehrt, W. T. Clark, P. Radivojac, M. W. Hahn, *PLOS Comput. Biol.* **7**, e1002073 (2011).
3. X. Chen, J. Zhang, *PLOS Comput. Biol.* **8**, e1002784 (2012).
4. L. Kapitzky *et al.*, *Mol. Syst. Biol.* **6**, 451 (2010).
5. C. J. Ryan *et al.*, *Mol. Cell* **46**, 691–704 (2012).
6. A. Frost *et al.*, *Cell* **149**, 1339–1352 (2012).
7. E. J. P. Douzery, E. A. Snell, E. Bapteste, F. Delsuc, H. Philippe, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15386–15391 (2004).
8. K. P. O'Brien, M. Remm, E. L. L. Sonnhammer, *Nucleic Acids Res.* **33**, D476–D480 (2005).
9. K. Dolinski, D. Botstein, *Annu. Rev. Genet.* **41**, 465–507 (2007).
10. M. E. Basson, M. Thorsness, J. Finer-Moore, R. M. Stroud, J. Rine, *Mol. Cell. Biol.* **8**, 3797–3808 (1988).
11. B. P. Cormack, M. Strubin, L. A. Stargell, K. Struhl, *Genes Dev.* **8**, 1335–1343 (1994).
12. M. J. Osborn, J. R. Miller, *Brief. Funct. Genomics Proteomics* **6**, 104–111 (2007).
13. A. H. Tong *et al.*, *Science* **294**, 2364–2368 (2001).
14. E. A. Winzeler *et al.*, *Science* **285**, 901–906 (1999).
15. The MGC Project TeamG. Temple *et al.*, *Genome Res.* **19**, 2324–2333 (2009).
16. J.-F. Rual *et al.*, *Nature* **437**, 1173–1178 (2005).
17. S. Mnaimneh *et al.*, *Cell* **118**, 31–44 (2004).
18. Z. Li *et al.*, *Nat. Biotechnol.* **29**, 361–367 (2011).
19. X. Pan *et al.*, *Mol. Cell* **16**, 487–496 (2004).
20. M. Kanehisa, S. Goto, *Nucleic Acids Res.* **28**, 27–30 (2000).
21. M. Groll *et al.*, *Nature* **386**, 463–471 (1997).
22. M. Groll *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10976–10983 (1999).
23. H. A. Orr, *Genetics* **139**, 1805–1813 (1995).
24. J. J. Welch, *Evolution* **58**, 1145–1156 (2004).
25. R. Guerois, J. E. Nielsen, L. Serrano, *J. Mol. Biol.* **320**, 369–387 (2002).
26. N. J. Marini, P. D. Thomas, J. Rine, *PLOS Genet.* **6**, e1000968 (2010).
27. S. M. Houten, H. R. Waterham, *Mol. Genet. Metab.* **72**, 273–276 (2001).

ACKNOWLEDGMENTS

We thank M. Minnix and A. Royall for assistance with cloning and assays, K. Drew for structural modeling assistance, M. Tsechansky for TANGO assistance, and C. Boone for providing the temperature-sensitive yeast strain collection. This work was supported by Cancer Prevention and Research Institute of Texas (CPRIT) research fellowships to A.H.K. and J.M.L; NIH grant R01 GM088344, Defense Threat Reduction Agency grant HDTRA1-12-C-0007, and NSF Science and Technology Center BEACON funds (DBI-0939454) to C.O.W.; and grants from the NIH, NSF, CPRIT, and the Welch foundation (F-1515) to E.M.M.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6237/921/suppl/DC1
Materials and Methods
Figs. S1 to S18
Tables S1 to S6
References (*28*–*55*)
Data S1 to S3

13 October 2014; accepted 21 April 2015
10.1126/science.aaa0769

Systematic humanization of yeast genes reveals conserved functions and genetic modularity

Aashiq H. Kachroo *et al.*

Science **348**, 921 (2015);
DOI: 10.1126/science.aaa0769

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 21, 2015):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/348/6237/921.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2015/05/20/348.6237.921.DC1.html

This article **cites 53 articles**, 24 of which can be accessed free:
http://www.sciencemag.org/content/348/6237/921.full.html#ref-list-1

This article appears in the following **subject collections:**
Genetics
http://www.sciencemag.org/cgi/collection/genetics

Downloaded from www.sciencemag.org on May 21, 2015

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2015 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.