Science Advances

www.advances.sciencemag.org/cgi/content/full/1/3/e1500248/DC1

Supplementary Materials for

A *Candida albicans* CRISPR system permits genetic engineering of essential genes and gene families

Valmik K. Vyas, M. Inmaculada Barrasa, Gerald R. Fink

Published 3 April 2015, *Sci. Adv.* **1**, e1500248 (2015)
DOI: 10.1126/sciadv.1500248

The PDF file includes:

CaCas9 protein sequence
Fig. S1. Recyclable Solo system vector pV1200 permits serial mutagenesis.
Fig. S2. *Candida* CRISPR-Duet system requires Cas9, sgRNA expression, and a mutagenic repair template.
Fig. S3. *Candida* CRISPR-Solo system requires a mutagenic repair template, but does not require selection for system components.
Fig. S4. *Candida* CRISPR permits the isolation of homozygous mutants at multiple loci, including *MtlA1*, *Mtla2*, *TPK2*, and *DCR1*.
Fig. S5. Mutation of *CDR1* and *CDR2* creates pleiotropic drug sensitivity.
Fig. S6. Mutation of *SNF1* in *Candida*.
Table S1. Plasmids used in this study.

Other Supplementary Material for this manuscript includes the following:
(available at www.advances.sciencemag.org/cgi/content/full/1/3/e1500248/DC1)

Supplementary Data Files (separate zip file)

CaCas9 protein sequence
MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR
RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK
LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL
SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI
GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD
QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR
RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM
TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFK
EDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ
KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNY
HHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE
QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDEGADPKKKRKVDPKKKRKVDPKKKRKVDYKDH
DGDYKDHDIDYKDDDDK

3x SV40 NLS
3x FLAG epitope

***CaCas9* DNA sequence**
ATGGATAAAAAGTATAGTATTGGTTTAGATATTGGTACTAACTCTGTGGGTTGGGCAGTTATCACCG
ACGAATATAAAGTTCCATCAAAGAAATTTAAGGTGTTAGGTAACACTGACAGACACTCAATAAAAA
AGAATCTTATCGGTGCTCTTTTGTTCGACTCCGGTGAAACTGCCGAGGCTACACGTTTAAAAAGAAC
AGCAAGAAGAAGATATACCCGTAGAAAAAATAGAATATGTTATTTACAAGAAATCTTTTCTAATGA
AATGGCTAAAGTTGATGATTCCTTTTTCCATAGATTGGAAGAGTCATTTTTGGTTGAAGAAGACAAA
AAGCATGAGAGACATCCAATCTTTGGGAATATAGTTGATGAAGTGGCTTACCATGAAAAATATCCT
ACCATTTATCATTTAAGAAAGAAATTGGTAGATTCAACTGATAAAGCTGACCTTAGATTAATCTATT
TAGCACTTGCCCATATGATTAAATTTAGAGGTCATTTTTTGATTGAAGGTGATTTGAACCCAGATAA
TTCTGACGTGGATAAATTATTTATTCAATTAGTCCAAACCTACAACCAATTATTTGAGGAAAATCCA
ATTAATGCTAGTGGTGTCGATGCCAAAGCTATATTATCAGCCAGATTATCAAAATCTAGACGTTTGG
AAAATTTGATTGCCCAATTGCCAGGAGAAAAAAAGAATGGATTATTTGGAAACTTGATCGCATTATC
ATTGGGTTTGACACCAAATTTTAAATCTAATTTTGATTTAGCTGAAGATGCTAAATTACAATTATCAA
AAGACACCTATGACGACGATTTGGACAATTTACTTGCTCAAATTGGTGATCAATATGCAGATTTGTT
CTTAGCTGCTAAAAACTTATCTGATGCTATTTTGTTGTCTGATATTTTGAGAGTGAACACAGAAATA
ACCAAAGCTCCATTATCAGCATCTATGATCAAACGTTATGATGAACACCATCAGGATTTGACTTTAT
TGAAAGCTTTGGTGAGACAACAATTGCCAGAGAAGTATAAAGAAATCTTTTTCGATCAATCTAAAA
ACGGGTATGCAGGTTATATTGATGGGGGTGCCTCCCAAGAGGAATTTTACAAATTTATAAAACCTAT
TTTAGAAAAGATGGATGGGACTGAGGAACTTTTGGTCAAATTGAACAGAGAAGATTTGTTACGTAA
ACAGAGAACTTTTGATAATGGTAGTATACCTCACCAAATTCATTTGGGTGAGTTGCATGCAATTTTA
AGAAGACAAGAAGATTTTTATCCATTTTTAAAAGATAATAGAGAAAAAATCGAGAAAATTTTAACC
TTTAGAATTCCATACTATGTTGGGCCTTTGGCTAGAGGTAATTCAAGATTTGCCTGGATGACACGTA
AATCAGAAGAAACTATTACCCCTTGGAATTTTGAAGAGGTTGTTGATAAAGGAGCATCAGCACAGA
GTTTTATTGAAAGAATGACCAATTTCGATAAAAACTTACCAAATGAAAAAGTTTTACCAAAACATTC
CTTGTTATACGAATATTTTACTGTTTACAATGAACTTACAAAGGTTAAATATGTTACTGAAGGTATGC
GTAAGCCAGCCTTTTTATCTGGAGAACAGAAAAAGGCAATAGTTGATTTATTGTTTAAAACAAATAG
AAAAGTTACTGTTAAACAATTAAAAGAAGATTACTTTAAGAAAATTGAATGTTTTGATTCAGTTGAA
ATCAGTGGTGTTGAAGACAGATTTAATGCTAGTTTAGGAACTTACCATGATTTACTTAAAATTATCA

AAGATAAAGATTTCTTGGATAACGAAGAAAATGAAGACATTTTAGAAGACATTGTTTTAACCTTAAC
TTTATTCGAAGATAGAGAGATGATTGAAGAACGTTTGAAGACTTATGCACATTTGTTTGACGATAAA
GTGATGAAACAGTTGAAAAGAAGACGTTATACTGGATGGGGTAGATTGTCTCGTAAATTGATCAAT
GGAATTAGAGATAAACAAAGTGGTAAAACTATCTTGGACTTTTTGAAATCTGACGGATTTGCTAATA
GAAATTTCATGCAATTGATCCACGACGATAGTTTGACATTTAAAGAAGACATCCAAAAGGCCCAAG
TGAGTGGGCAAGGTGATTCATTACATGAACATATTGCAAATTTAGCCGGATCTCCTGCTATTAAGAA
AGGGATATTACAAACTGTTAAAGTTGTGGATGAATTAGTGAAAGTAATGGGAAGACATAAACCTGA
AAACATTGTCATTGAGATGGCAAGAGAAAATCAAACTACACAAAAAGGACAGAAAAATAGTAGAG
AACGTATGAAAAGAATAGAAGAGGGTATTAAAGAATTGGGTAGTCAAATATTGAAAGAACACCCA
GTGGAAAATACCCAGTTGCAAAATGAAAAATTATATCTTTACTACCTTCAAAATGGACGTGATATGT
ATGTTGATCAGGAATTAGATATAAATAGACTTTCAGATTATGATGTAGATCATATAGTTCCACAATC
TTTCTTGAAAGATGATTCCATAGACAATAAAGTATTAACTAGAAGTGATAAAAATAGAGGTAAAAG
TGATAATGTCCCAAGTGAGGAAGTCGTCAAAAAGATGAAAAATTACTGGCGTCAACTTTTGAATGC
TAAATTAATTACTCAAAGAAAATTTGATAATTTGACTAAAGCAGAAAGAGGTGGGCTTTCTGAATTA
GATAAAGCCGGGTTCATTAAAAGACAATTGGTCGAAACTAGACAAATTACTAAACATGTTGCCCAA
ATTTTAGATTCCCGTATGAACACTAAGTATGACGAAAATGATAAGTTAATACGTGAGGTTAAAGTCA
TTACTTTAAAATCAAAACTTGTCTCTGATTTCAGAAAGGATTTCCAATTCTATAAAGTTAGAGAAAT
TAATAATTATCATCATGCTCATGATGCATATTTGAATGCTGTAGTTGGAACTGCTTTAATCAAGAAA
TACCCTAAATTAGAATCTGAATTTGTATATGGTGATTACAAAGTCTATGATGTTAGAAAGATGATTG
CTAAATCAGAACAAGAAATTGGTAAAGCTACAGCTAAATACTTCTTTTACTCTAACATTATGAATTT
CTTTAAAACAGAAATTACTTTGGCAAACGGTGAAATTAGAAAAAGACCTCTTATTGAAACAAATGG
TGAGACTGGAGAGATAGTTTGGGACAAAGGGCGTGATTTCGCTACTGTTAGAAAAGTTTTATCAATG
CCACAAGTTAACATTGTAAAGAAAACAGAGGTTCAAACTGGTGGTTTCTCAAAAGAAAGTATTTTG
CCTAAAAGAAATAGTGATAAATTGATTGCCAGAAAAAAGGATTGGGATCCAAAGAAATATGGTGGT
TTCGACTCACCAACCGTAGCCTATTCTGTTTTGGTTGTGGCAAAGGTTGAAAAGGGTAAAAGTAAAA
AGCTTAAATCAGTAAAAGAACTTTTGGGTATTACAATAATGGAAAGAAGTTCCTTTGAAAAGAACC
CTATTGATTTTTTGGAAGCTAAAGGTTATAAGGAAGTAAAGAAGGACTTAATAATCAAATTGCCTAA
ATATTCTTTATTTGAATTAGAAAATGGGAGAAAAAGAATGTTGGCTTCTGCTGGAGAATTGCAAAAG
GGTAATGAATTAGCATTGCCTTCCAAATATGTTAACTTCTTGTATTTAGCTTCACACTATGAAAAGTT
GAAAGGGTCACCAGAAGATAACGAGCAAAAACAATTATTTGTTGAACAACACAAACACTACTTAGA
TGAGATTATAGAACAAATTAGTGAATTCAGTAAAAGAGTGATATTAGCTGATGCAAATTTAGATAA
AGTTTTGTCAGCCTATAACAAACATAGAGATAAGCCAATTAGAGAACAAGCAGAAAACATTATTCA
CTTATTTACCCTTACCAATTTAGGAGCACCTGCTGCTTTCAAGTATTTTGATACAACAATTGATCGTA
AAAGATATACCTCAACAAAAGAAGTCTTAGACGCCACCTTAATTCATCAATCAATCACTGGATTGTA
TGAGACAAGAATTGATTTGTCTCAATTGGGTGGTGATGAAGGGGCTGATCCTAAGAAGAAAAGAAA
AGTTGATCCAAAGAAAAAGCGTAAGGTGGATCCTAAGAAAAAGAGAAAGGTTgactacaaagaccatgacggt
gattataaagatcatgacatcgactacaaggatgacgatgacaagTGATAA

3x SV40 NLS
3x FLAG epitope
2x stop codons

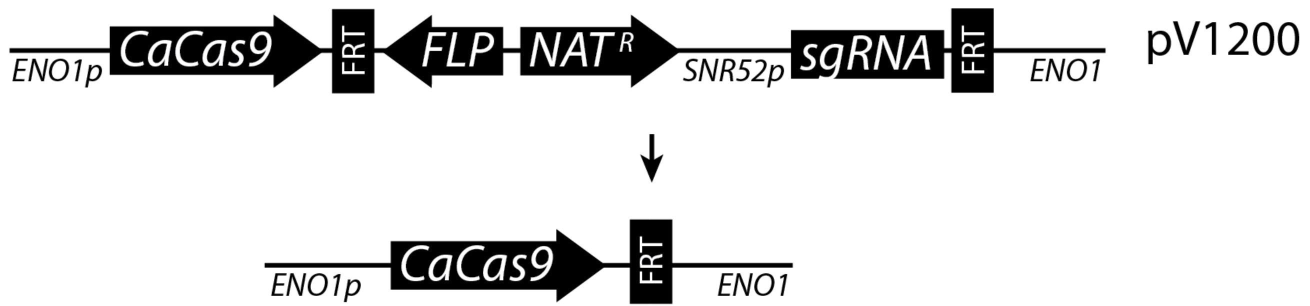


Figure S1. Recyclable Solo system vector pV1200 permits serial mutagenesis. The pV1200 Solo system vector is identical to the Solo system vector pV1093, except that it contains the *Nat*^{R}*-FLP* and *SNR52p-sgRNA* cassette flanked by FRT sites, and an inducible Flippase under the control of the *SAP2* promoter. Induction of Flippase causes excision of the *Nat*^{R}*-FLP-SNR52-sgRNA* cassette (bottom), leaving a Nat sensitive strain that can be mutagenized with another sgRNA expression cassette.

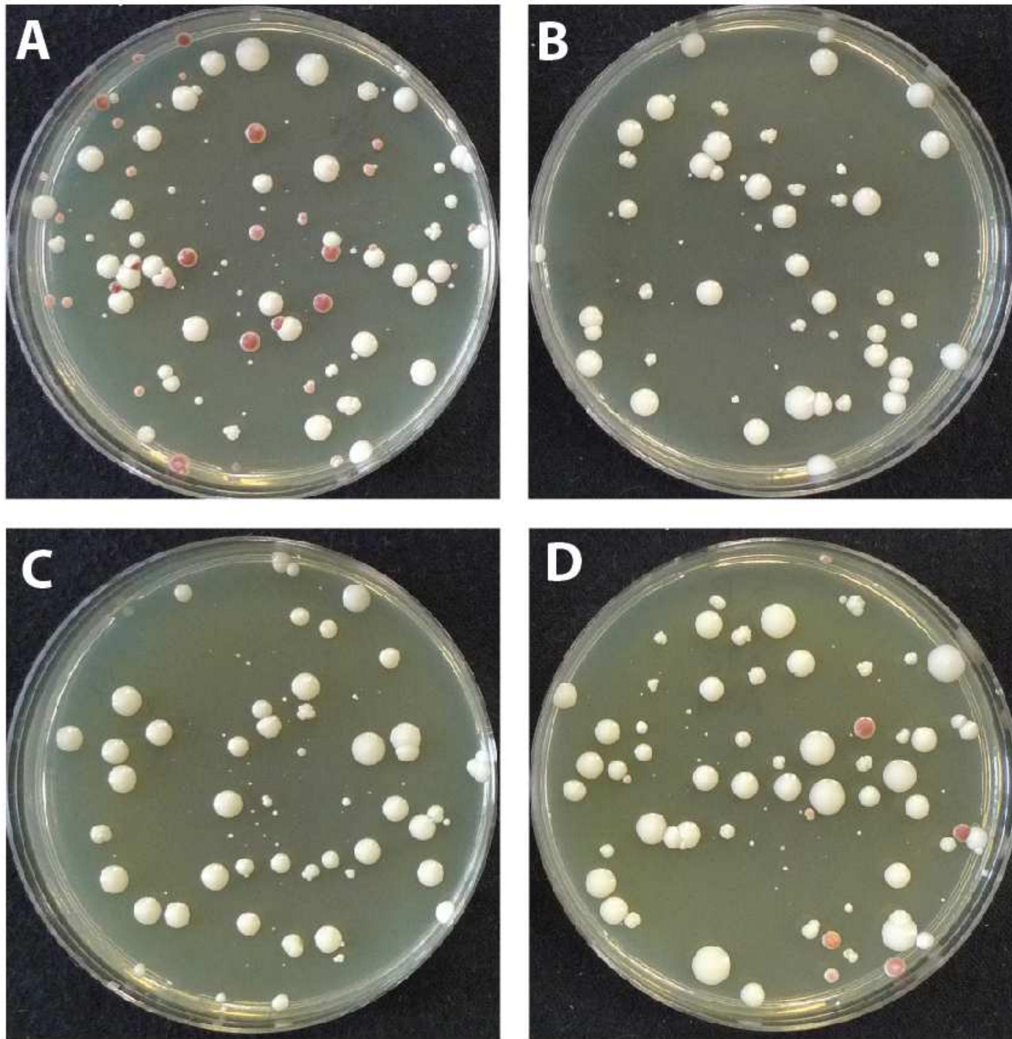


Figure S2. ***Candida*** **CRISPR-Duet system requires Cas9, sgRNA expression, and a mutagenic repair template.** Strain VY959 (**A**+**B**), which contains the integrated Cas9 from the Duet system, was transformed with pV1010 (Duet sg*ADE2* expression plasmid) without (**B**) or with (**A**) a mutagenic repair template, and plated on YPD+Nat. Strain SC5314 (**C**+**D**) was transformed with pV1010 with a repair template without (**C**) or with (**D**) Cas9 expression plasmid pV1025.

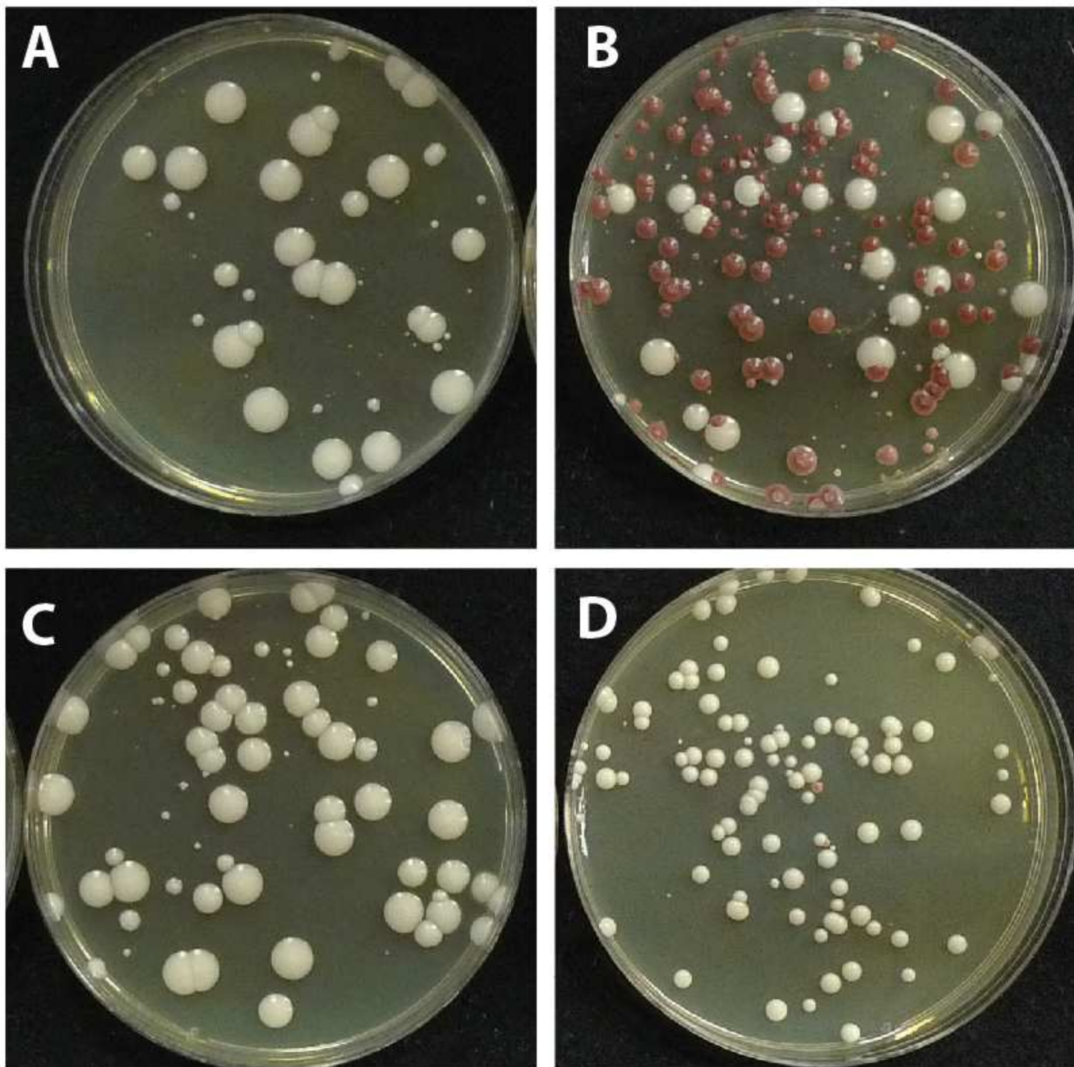
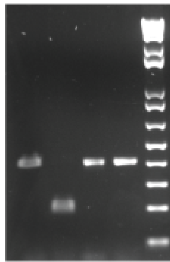
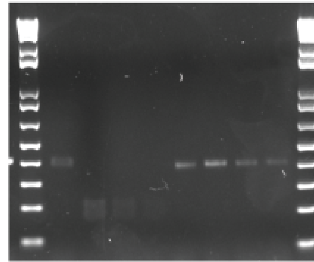
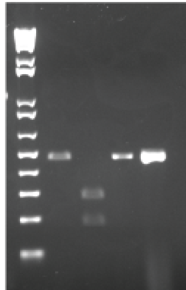


Figure S3. ***Candida*** **CRISPR-Solo system requires a mutagenic repair template, but does not require selection for system components.** Strain SC5314 was transformed with pV1081 (Solo system for *ADE2*) without (**A**) or with a mutagenic template containing the guide sequence (**B**) or 250-bp downstream (**C**), and plated on YPD+Nat. Dilution from (**B**) was plated to non-selective YPD plates (**D**).







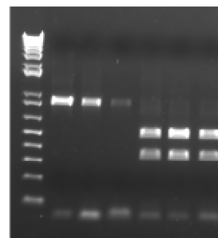


Figure S4. *Candida* CRISPR permits the isolation of homozygous mutants at multiple loci, including *MtlA1, Mtlα2, TPK2,* and *DCR1*. PCR genotyping of indicated genes is shown, and numbers listed are base pair positions with respect to the ATG codon.

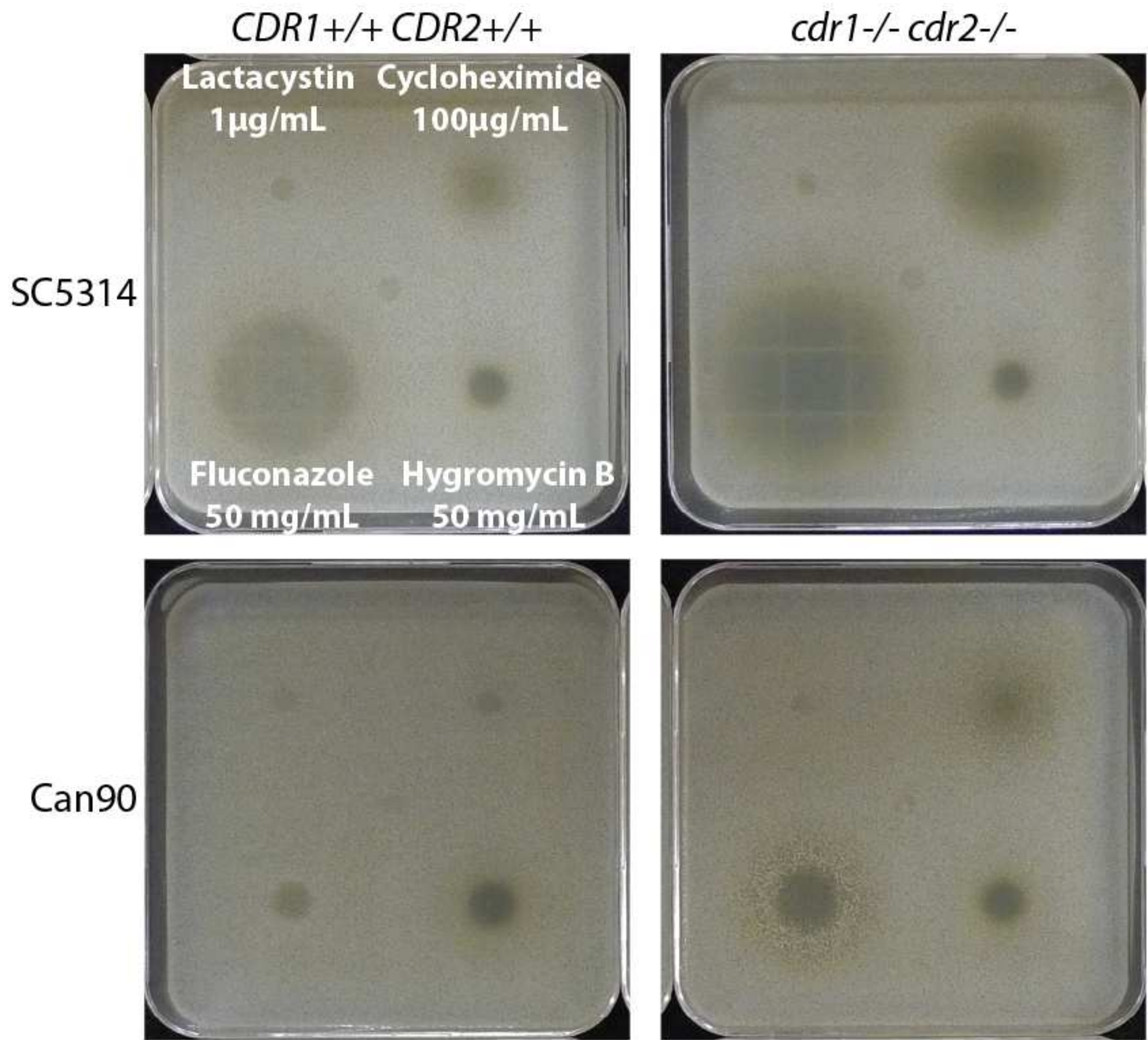


Figure S5. Mutation of *CDR1* and *CDR2* creates pleotropic drug sensitivity. Three microliters of the indicated drugs were spotted atop YPD plates containing the indicated strain. Plates were allowed to grow overnight and photographed.

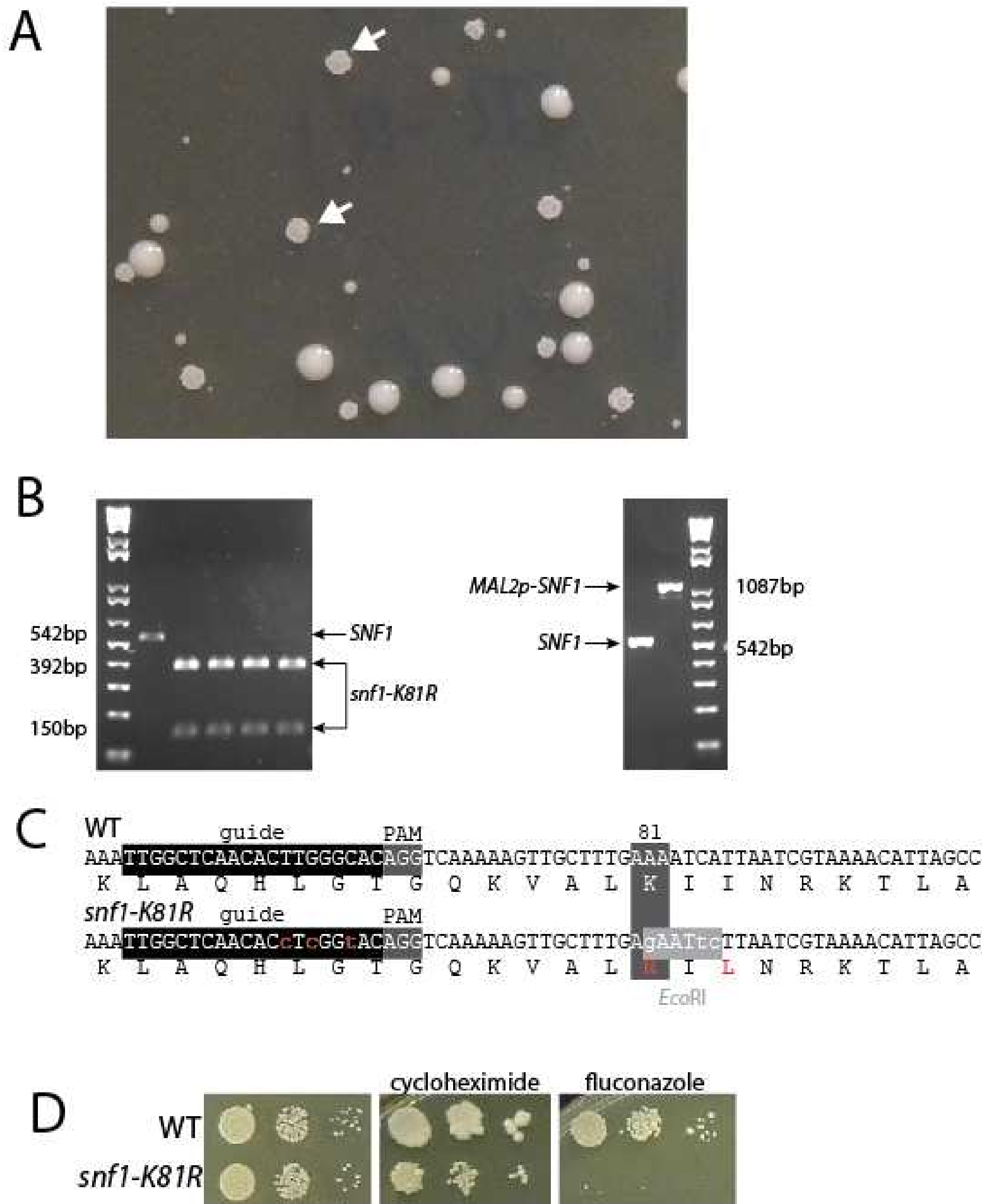


Figure S6. Mutation of *SNF1* in *Candida*. A. Unusual colony morphology of *snf1-K81R* transformants. Wrinkly colonies (two examples are marked with arrows) contain the K81R mutation, while smooth colonies are WT. **B**. PCR confirmation of homozygous *SNF1* mutation. Mutation at position K81R introduces an *Eco*RI site not found in the WT locus (left) and insertion of *MAL2p* at *SNF1* increases size of PCR amplification with *SNF1* primers (right). **C.** Sequence of WT and *snf1-K81R* alleles. Silent mutations were introduced into targeting region to prevent further cleavage. **D.**

Strains of the indicated genotype were grown in YPD alone or with cycloheximide (400µg/ml) or fluconazole (1µg/ml).

Table 1: Plasmids used in this study	
pV1025	Duet system CaCas9 expression vector, contains Nat^R/FLP cassette, and targeting arms for the *ENO1* locus. The *ENO1p* is used to drive CaCas9 expression.
pV1090	Duet system sgRNA entry expression vector, contains Nat^R gene and the *SNR52* promoter from Candida albicans driving expression of sgRNA that binds/targets Cas9, and targeting arms to direct integration to *RP10.*
pV1093	Solo system CaCas9/sgRNA entry expression vector, contains Nat^R gene, and 2kb targeting arms for the upstream and downstream of *ENO1* coding region. *ENO1p* drives CaCas9 expression as above.
pV1081	Solo system vector to target mutagenesis of *ADE2*
pV1086	Solo system vector to target mutagenesis of *CDR1 and CDR2*
pV1102	Solo system vector to target mutagenesis of *URA3*
pV1107	Solo system vector to target mutagenesis of *RAS1*
pV1123	Solo system vector to target mutagenesis of *MtlA1*
pV1126	Solo system vector to target mutagenesis of *MtlAlpha2*
pV1147	Solo system vector to target mutagenesis of *TPK2*
pV1129	Solo system vector to target mutagenesis of *DCR1,* first position
pV1132	Solo system vector to target mutagenesis of *DCR1,* second position
pV1138	Solo system vector to target mutagenesis of *SNF1* proximal to K81
pV1144	Solo system vector to target mutagenesis of *SNF1* promoter
pV1200	Solo system CaCas9/sgRNA entry expression vector, contains Nat^R gene, and 2kb targeting arms for the upstream and downstream of *ENO1* coding region. *ENO1p* drives CaCas9 expression as above. The Nat^R gene and *SNR52p-sgRNA* cassette is flanked by FRT sites, which mediate recombination when *FLP* expression is induced.

Oligonucleotide sequences used in this study

sgRNA cloning Primers

sgADE2 top	atttgCAACAATCATACGACCTAATg
sgADE2 bottom	AAAACattaggtcgtatgattgttgc
sgURA3 top	atttgAGTTTCTGCTCTCTCACTATg
sgURA3 bottom	AAAACatagtgagagagcagaaactc
sgRAS1 top	atttgAAATTAGTTGTTGTTGGAGGG
sgRAS1 bottom	AAAACCCTCCAACAACAACTAATTTc
sgMtlA1 top	atttgATATAAGAATGAAGACAACGg
sgMtlA1 bottom	aaaacCGTTGTCTTCATTCTTATATc
sgMtlAlpha2 top	atttgACAAGACATGAATTCACATCG
sgMtlAlpha2 bottom	AAAACGATGTGAATTCATGTCTTGTc
sgSnf1p top	atttgATATAATGTGTATTACTTCTG
sgSnf1p bottom	AAAACAGAAGTAATACACATTATATc
sgSnf1-1 top	atttgTTGGCTCAACACTTGGGCACG
sgSnf1-1 bottom	AAAACGTGCCCAAGTGTTGAGCCAAc
sgDcr1-1 top	atttgATAGCAGAAACTGCCAACAAg
sgDcr1-1 bottom	aaaacTTGTTGGCAGTTTCTGCTATc
sgDcr1-2 top	atttgTTATGAGTTACATCAACAACg
sgDcr1-2 bottom	aaaacGTTGTTGATGTAACTCATAAc
sgTpk2 top	atttgGGGTGAACTATTTGTTCGCCG
sgTpk2 bottom	AAAACGGCGAACAAATAGTTCACCCc

PCR/Sequencing Primers

ADE2-fwd	aacaccccccaccaaaaagaatc
ADE2-rev	acaagtcatcgactgtgttgg
CDR1-fwd	AAAACATTCAGAATTTAGCCAG
CDR2-fwd	atagaaatttaagagcttacgg
CDR12-rev	aggttgccatataaacactagcc
URA3-fwd	tttgttcttcaatgatgatttcaacc
URA3-rev	cataaattgatgtttacgtgaaagttc
RAS1-fwd	TCAATTGACTAGATATAAACTCTTC
RAS1-rev	TCCATCTTCATAACTAACTTGTCTT
MatA1-fwd	TTCAATAGTTTTTTTCTGCGTATTGTG
MtlA1-rev	TCGATCCAGCAATGGAAGATAGCTT
MtlAlpha2-fwd	CTTAGTCTAACTTTATAGTTGTC
MtlAlpha2-rev	ATTCTTTCTAATAACATTTCATGCAA
Snf1-fwd	TGTCATTCCGTTTCTCCTTCTA
Snf1-rev	GCAAATTCAATAACCATAATG
DCR1-fwd	GGTATATTGCACACGACCATAGTGCGAA
DCR1-rev	TCACTTATTTTGACTTCATC
Tpk2-fwd	TTAAAGAAACTTCACATCACCAA
Tpk2-rev	ACTTTGATAGCATAATATCTAC

Repair Templates for mutagenesis

ADE2-NT2-top	taatggatagcaaaactgttggtattttaggaggttaatgattaggtcgtatgattgttgaagcag
ADE2-NT2-bottom	cggtcttgatattcaatctatgtgctgcttcaacaatcatacgacctaat
ADE2-NT1-top	ttgatgttgatgctttaatcaaagttcaagagaaattAACtaaagttgaaatatatccattacTAC CTGAAAC
ADE2-NT1-bottom	tatcttgaatcaatcttatggtttcaggtaatggatatatttcaacttta
CDR12-top	ccaggtgaacttactgtKgttttggggagacccggtgctTAAGaaTTCttgttccacatt
CDR12-bottom	tgtggaaaccataagtgttaacagcaatggtctttaacaatgtggaacaaGAAttCTTAa
URA3-top	aaatagcaaacaaaagatatgacagtcaacactTAATAATatagtgagagagcagaaact
URA3-bottom	aaataatcgttgtgctactggtgaggcatgagtttctgctctctcactat
RAS1-V13-top	ATATCCACACATATACATACCATGTTGAGAGAATATAAATTAGTTGTTGTTGGAGGTGtT
RAS1-V13-bottom	AATCAATTGAATGGTTAAAGCGGATTTACCAACACCAaCACCTCCAACAACAACTAATTT
RAS1-TAA13-top	ATATCCACACATATACATACCATGTTGAGAGAATATAAATTAGTTGTTGTTGGAGGTtaa
RAS1-TAA13-bottom	AATCAATTGAATGGTTAAAGCGGATTTACCAACACCgaattcttaACCTCCAACAACAAC
MtlA1-top	TTTAAAAAGTGTAGAGAAACTAGTTCAAGCAACATCAGTATATAAGAATGAAGACAACGA
MtlA1-bottom	TGCCTCTCACGCTTCAATTGTAAGAATATTTgaattcatTCGTTGTCTTCATTCTTATAT
MtlAlpha2-top	ACAACACTAACTCGGTACTCAAGTTATACTCACATCAATAACAAGACATGAATTCACATC
MtlAlpha2-bottom	GCAAGCGTTGATTTATTTCAAAGAGTGCCTCggatccttaaAGATGTGAATTCATGTCTT
Snf1-Mal-PCR-top-fwd	TTCACAGAGTGATTATCTGAGTCGTTCATACACCCAAGAAGTTTGATATTTTTGTCTAGT
Snf1-Mal-PCR-bottom-rev	TGACATCTTTAACTCTATGTTATTATATAATGTGTATTACCATTGTAGTTGATTATTAGT
Snf1K81R-top	CTCAAGACATTAGGTGAAGGGTCATTTGGTAAAGTGAAATTGGCTCAACACcTcGGtACAGGTCAA AAAGTTGCTTTGAgAAT
Snf1K81R-bottom	TAAATATGAAATCTCTCTTTCAACACGACCCTGCATGTCgcttTTtGCTAATGTTTTACGATTAAT aATTcTCAAAGCAACTTT
Snf1K81R-EcoR1-bottom	TAAATATGAAATCTCTCTTTCAACACGACCCTGCATGTCgcttTTtGCTAATGTTTTACGATTAAg aATTcTCAAAGCAACTTT
DCR1-1-top	TTTTCTCAAAAAAATCTAGCAGCACAAAATATAGCAGAAACTGCCAACAAAtaagaattc
DCR1-1-bottom	GTTGACTGGTAGATGTCCAGTTGTTGATGTAACTCATAAAgaattcttaTTTGTTGGCA
DCR1-2-top	TAGCAGCACAAAATATAGCAGAAACTGCCAACAAAGGGTTTATGAGTTACATCAACAACT
DCR1-2-bottom	ACTTTATTATCTTCTTGTTGACTGGTAGATGTgaattcttAGTTGTTGATGTAACTCATA
Tpk2-top	ACAATTTCAACAACCGCAGCAACAACTTTATtaAgaattcGGCGAACAAATAGTTCACCC
Tpk2-bottom	TGTTACATTTGTAGTATTTTGTCCAGTTTGGGCTGCAGCAGGGTGAACTATTTGTTCGCC

CDR1/2 guide sequence

GTTTTGGGGAGACCCGGTGC

Supplementary Data Files (separate zip file)
The Supplementary Data files are found in the zip file, 1500248_Supplementary Data Files.zip

The file consists of many tables that contain mappings of potential CRISPR target sites in the diploid assembly to the Candida albicans genome, retrieved from the Candida Genome Database (candidagenome.org) in December 2014. The files include:

GenesDontHaveUniqTargets.txt
GenesDontHaveUniqTargetsConsideringThe12ntSubs.txt
Targ.NoTs.subs12nt.HitsGenesOnly.Hits1Gene1Alle.3letterName.xlsx
Targ.NoTs.subs12nt.HitsGenesOnly.Hits1Gene2Alle.3letterName.xlsx
Targ.NoTs.subs12nt.HitsGenesOnly.Hits1Gene2AlleInMoreThan1Place.3letterName.xlsx
Targ.NoTs.subs12nt.HitsGenesOnly.MoreThan1Gene.OneRowPerGene.3letterName.xlsx
Targ.NoTs.subs13nt.HitsGenesOnly.Hits1Gene1Alle.distances.3letterName.xlsx
Targ.NoTs.subs13nt.HitsGenesOnly.Hits1Gene2Alle.distances.3letterName.xlsx
Targ.NoTs.subs13nt.HitsGenesOnly.Hits1Gene2AlleInMoreThan1Place.distances.3letterName.xlsx
Targ.NoTs.subs13nt.HitsGenesOnly.MoreThan1Gene.OneRowPerGene.3letterName.xlsx

Target sites with a string of 6 or more T's were excluded, as they would cause premature termination from the RNA Polymerase III promoter used in our system.

Searches for reduced stringency targeting were done using either 12 or 13 nucleotides most proximal to the PAM sequence, and the results are stored in separate files.

All of the mappings here target genes only - those with potential targeting to intergenic regions were excluded.

The mappings are grouped as follows:

Hits1Gene1Allele = Sites listed here are allele specific, and have no other sites in the genome
Hits1Gene2Allele = Sites listed here target both alleles of a given gene, and nowhere else.
Hits1Gene2AlleleInMoreThan1Place = Sites listed here target both genes, but may map to more than one position within each gene (e.g., genes with internal repeats that are unique)
MoreThan1Gene = Sites listed here hit multiple genes in the indicated position.

A small number of genes are not uniquely targetable by the 12 or 13 nucleotide criteria, and they are listed in the indicated files.