'Calling Cards' method for high-throughput identification of targets of yeast DNA-binding proteins

Haoyi Wang, Michael E Heinz, Seth D Crosby, Mark Johnston & Robi D Mitra

Department of Genetics, Washington University, School of Medicine, Genome Sequencing Center, Campus Box 8510, Room 5401, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA. Correspondence should be addressed to R.D.M. (rmitra@genetics.wustl.edu).

Published online 11 September 2008; doi:10.1038/nprot.2008.148

We present a protocol for a novel method for identifying the targets of DNA-binding proteins in the genome of the yeast *Saccharomyces cerevisiae*. This is accomplished by engineering a DNA-binding protein so that it leaves behind in the genome a permanent mark—a 'calling card'—that provides a record of that protein's visit to that region of the genome. The calling card is the yeast Ty5 retrotransposon, whose integrase interacts with the Sir4 protein. If Sir4 is fused to a DNA-binding protein, it recruits the Ty5 integrase, which directs insertion of a Ty5 calling card into the genome. The calling card along with the flanking genomic DNA is harvested by inverse PCR and its genomic location is determined by hybridization of the product to a DNA microarray. This method provides a straightforward alternative to the 'ChIP-chip' method for determining the targets of DNA-binding proteins. This protocol takes ~2 weeks to complete.

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols

INTRODUCTION

Transcription factors (TFs) bind to specific sites in the genome and control gene transcription. Identification of the genomic sites bound by all DNA-binding proteins will provide a detailed map of the transcriptional networks that direct different cellular processes and provide a framework for understanding how a cell controls global patterns of gene expression. Here we describe the 'calling card' method, a tool to provide this information.

Overview of the calling card method

The calling card method exploits the Ty5 retrotransposon of bakers' yeast. Ty5 mRNA is converted by reverse transcriptase into a double-stranded cDNA that the Ty5 integrase carries to the nucleus and inserts into the genome[1]. The Ty5 integrase interacts with the Sir4 heterochromatin protein[2]. Therefore, any DNA-binding protein can be made to recruit the Ty5 integrase by attaching to it the fragment of the Sir4 protein that interacts with the integrase[3]. Consequently, the engineered DNA-binding protein directs insertion of Ty5 into DNA near to where it is bound, leaving behind a permanent mark—a 'calling card'—of its visit to that region of the genome. We have exploited this property of Ty5 to develop a method for identifying the genomic targets of DNA-binding proteins.

The TF–Sir4 fusions are made by joining the TF of interest to a fragment of Sir4 (amino acids 951–1,200) that includes the Ty5 integrase-interacting domain[2,3]. We have been fusing the Sir4 fragment to the C terminus of the TFs, but it may be preferable to fuse it to the N terminus in some cases (e.g., if the DNA-binding domain of the TF is near the C terminus). Based on our experiments to date, no linker is necessary between TF and Sir4 protein. The TF–*SIR4* fusions can be constructed in yeast by the 'gap repair' method[4,5]. See REAGENT SETUP for details on primer design for obtaining TF coding sequence DNA by PCR.

The 'calling card' protocol, summarized in **Figures 1** and **2**, can be divided into five stages: (1) construction of a yeast strain carrying a plasmid encoding the desired TF–Sir4 chimera and a plasmid carrying Ty5, (2) induction of Ty5 transposition, (3) selection of cells that have undergone transposition of Ty5, (4) recovery of the Ty5 calling cards from genomic DNA by inverse PCR and (5) identification of the flanking genomic DNA sequence by hybridization of the inverse PCR product to a DNA microarray.

All the experiments should be done in a *sir4* deletion strain (e.g., YM7635), otherwise wild-type Sir4 protein will compete with TF–Sir4 for binding to Ty5 integrase, causing transposition into telomeres. There are three controls that should be used in any calling card experiment. First, one should analyze Ty5 transposition

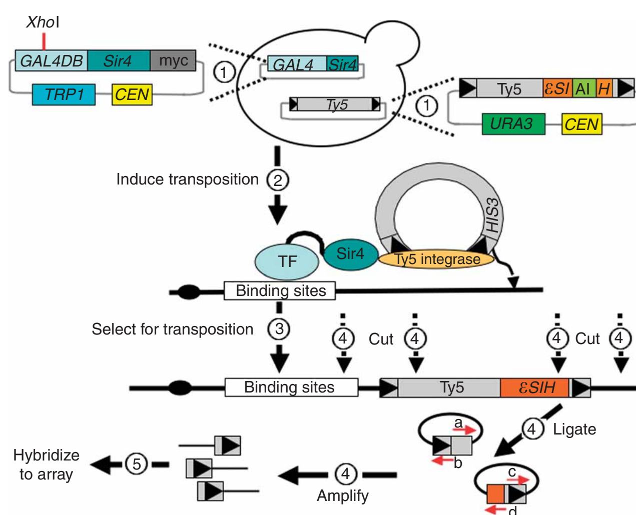

Figure 1 | The five stages of the 'calling cards' protocol. (1) construction of a yeast strain carrying a plasmid encoding the desired transcription factor (TF)–Sir4 fusion and a plasmid carrying Ty5. A *HIS3* marker is inserted into Ty5 in the opposite direction, within which lies an artificial intron (AI). The two black triangles at the ends of Ty5 represent the long terminal repeats (LTRs) of Ty5. (2) Induction of Ty5 transposition. (3) Selection of cells that have undergone transposition of Ty5. (4) Recovery of the Ty5 calling cards from genomic DNA by inverse PCR. Primer a is OM6458; primer b is OM6609; primer c is OM6610; primer d is OM6456. (5) Identification of the flanking genomic DNA sequence by hybridization of the inverse PCR product to a DNA microarray. Modified from Figure 1 in Wang *et al.*[8].

in the *sir4* deletion working strain, without any TF–Sir4 fusion construct (**Box 1**). This controls for background transposition that is not directed by the TF–Sir4 chimera. We have found the patterns of transposition to be significantly different between strains without the TF–Sir4 construct and strains expressing a TF–Sir4 fusion protein. A similar pattern of transposition in both strains is a clear indication that something is wrong with the TF–Sir4 construct. Second, to control for the variation in hybridization efficiency across different probes on the microarray, we label genomic DNA and use this as a hybridization control. The inverse PCR samples are labeled with cy5 and the genomic DNA is labeled with cy3. Both labeled samples are hybridized to the same microarray. The intensity values in the control channel (the green, or cy3 channel) are used to estimate the hybridization efficiency of each probe, which allows us to accurately quantify the amount of DNA hybridized in the experimental channel (the red, or cy5 channel). Finally, as a positive control, it is useful to analyze Ty5 transposition in yeast expressing a Gcn4–Sir4 fusion protein. This control, which only needs to be included the first time a calling card experiment is performed, is important for the analysis of the microarray hybridization because it determines the intensity cutoff that separates transposition events from hybridization noise.

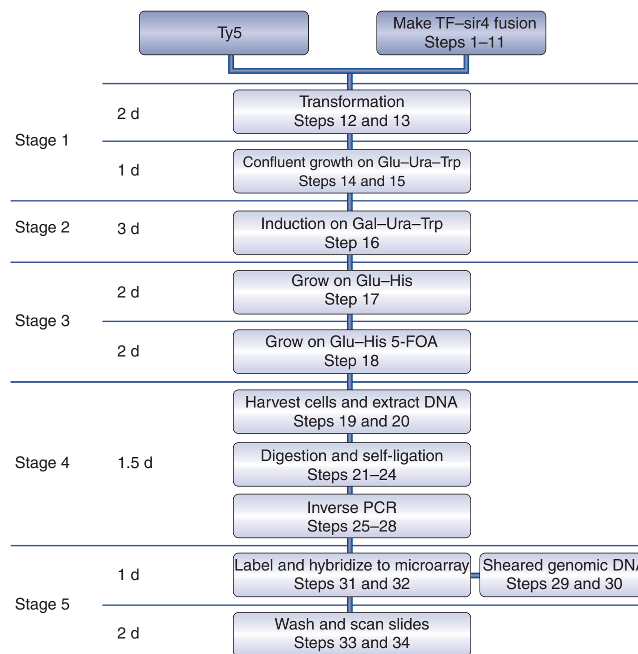

Figure 2 | A sample timing for the 'calling cards' protocol.

Advantages of the calling card method

We believe this technology will prove useful for the study of DNA-binding proteins because it is relatively easy to employ, and is in principle orthogonal to the ChIP-chip method[6,7]. Even if the calling card technology does not prove to be a substitute for the ChIP-chip method, it is likely to complement that well-established method because it can identify targets of proteins that may be refractory to analysis by chromatin IP and can be used to verify results obtained with the ChIP-chip method[8]. In addition, there are opportunities for multiplexing the calling card technology (using DNA barcodes), offering the possibility of identifying the targets of many DNA-binding proteins in a single experiment[8]. In this protocol, we focus on mapping genome-wide binding of a single TF, but this procedure can easily be extended using modifications detailed in our earlier article[8] to determine all TFs that bind to a single promoter. We are in the process of coupling calling card technology with Illumina 1G sequencing to analyze the genome-wide binding of multiple TFs in a single experiment.

Limitations of the calling card method

There are several limitations of this method in its current state: first, the transposition efficiency of Ty5 is fairly low ($\sim 10^{-5}$), which makes it difficult to sample more than a few thousand transposition events. Second, each TF is driven by the *ADH1* promoter, so its expression level is not native. Third, expression of the Ty5 calling card from the *GAL1* promoter limits the conditions that can be tested. Finally, Ty5 transposition is influenced by host factors[9], so the implementation of calling cards in an organism other than *Saccharomyces cerevisiae* will probably require the use of a different transposon.

Other methods for identifying target sites of DNA-binding proteins

ChIP-based methods. The 'ChIP-chip' method combines chromatin immunoprecipitation (ChIP) with DNA microarrays (chip): DNA is coprecipitated with a DNA-binding protein by ChIP, and then identified by hybridization to a DNA microarray[6,7]. ChIP is also now being combined with 'next generation' DNA sequencing[10,11]. ChIP-chip has been successfully applied to map the target genes of TFs in yeast[12,13] and other organisms[14,15]. The related ChIC and ChEC methods are tailored to the analysis of insoluble proteins, such as the scaffolding components of chromatin[16]. These

BOX 1 | EXPERIMENTAL CONTROL: Ty5 TRANSPOSITION WITHOUT TF–Sir4 (PERFORM IN PARALLEL TO THE MAIN PROTOCOL STEPS 12–16)

1. Transform the plasmid pBM5218 containing the Ty5 transposon (with *URA3* marker) into yeast strain YM7635 as described in **Box 2**. Use 0.1–0.5 µg of plasmid DNA for transformation. Select transformants on synthetic complete (SC) Glucose–Ura plates.
2. After 2-d incubation at 30 °C, multiple yeast colonies should be observed on SC Glucose–Ura plate. From each plate, pick one colony and culture overnight in 5 ml SC Glucose–Ura media at 30 °C.
3. Once the culture reaches an OD_{600} of 1 or higher, plate 500 µl of cells on each of ten SC Glucose–Ura plates.
4. Grow at 30 °C for 1 d until a confluent lawn is formed.
5. Replica plate the cells onto SC Galactose–Ura plates to induce Ty5 transposition. Keep plates at room temperature (22–25 °C) for 3 d.
6. Continue with the main protocol from Step 17.

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols

methods are similar to ChIP in that DNA-binding proteins are crosslinked to DNA, but they employ a micrococcal nuclease that is tethered to an antibody (ChIC) or the DNA-binding protein itself (ChEC), to introduce double-stranded breaks in unbound DNA.

The ChIP-based methods are powerful because they are highly flexible—they can be used to analyze a wide variety of DNA-binding proteins in a number of model systems. One weakness is that the results of ChIP-type experiments often depend on the quality of the antibody employed, although this can be somewhat alleviated by expressing DNA-binding proteins with peptide tags. Also, some DNA-binding proteins appear to be recalcitrant to ChIP-chip and related methods[13].

Yeast-1 hybrid. A one-hybrid screen can also be used to identify the TFs that bind to a specific genomic locus[17]. In this method, a query sequence is cloned in front of a reporter gene, and a library of TF-activation domain fusion constructs are screened[18]. This method has been used to reveal the architecture of regulatory networks in *C. elegans*[19,20]. A strength of this method is that it is easily automatable, allowing for high-throughput analysis of many loci and proteins. It has the disadvantage that TF binding is not queried at the native locus.

DamID. Another method for the identification of DNA loci bound by TFs is DamID[21,22]. In this method, a DNA adenine methyltransferase (Dam) is fused to a TF, which targets DNA methylation to adenines that are close to binding sites. This method can be used to analyze proteins that are resistant to ChIP-chip. One possible weakness is that nonspecific methylation often occurs, although this can be addressed with the appropriate controls.

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols

MATERIALS

REAGENTS

- Diploid yeast strain with sir4 deletion, YM7635 (MATa /MATalpha his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 ura3Δ0/ura3Δ0 met15Δ0/MET15 lys2Δ0/LYS2 sir4::Kan/sir4::Kan trp1::Hyg/ trp1::Hyg)
- Plasmid pBM4607 (contains the Gal4DB–*sir4* fusion with *Trp1* as the selectable marker) (sequence has been submitted to addgene.org, Plasmid 18795)
- Plasmid pBM5218 (encodes the Ty5 transposon with *URA3* as the selectable marker) (sequence has been submitted to http://addgene.org/, Plasmid 18796)
- Restriction enzymes: *Xho*I, *Hin*P1I, *Hpa*II and *Taq*I (NEB)
- QIAquick Gel Extraction Kit (Qiagen, cat. no. 28704)
- Phusion DNA polymerase (with manufacturer's buffers; NEB, cat. no. F-530S)
- 10 mM dNTP mix (Roche, cat. no. 12779120)
- Yeast peptone dextrose (YPD) (see REAGENT SETUP)
- Synthetic complete (SC) Glucose–Trp; used to select for the plasmid pBM4607, which contains a *TRP1* marker (see REAGENT SETUP)
- SC Glucose–Ura; used to select for the plasmid pBM5218, which contains a *URA3* marker (see REAGENT SETUP)
- SC Glucose–Ura–Trp; used to select for pBM5218 and pBM4607 (see REAGENT SETUP)
- SC Galactose–Ura; used to select for pBM5218, and to activate the GAL1-10 promoter (see REAGENT SETUP)
- SC Galactose–Ura–Trp; used to select for pBM5218 and pBM4607, and to activate the GAL1-10 promoter (see REAGENT SETUP)
- SC Glucose–His; used to select for cells with a Ty5 transposition event (see REAGENT SETUP)
- SC Glucose–His 5-fluoroorotic acid (5-FOA) used to select for cells with a Ty5 transposition event and to select against the Ty5 donor plasmid pBM5218 (see REAGENT SETUP)
- Yeast lysis buffer (see REAGENT SETUP)
- Phenol/chloroform/iso-amyl alcohol (25:24:1; Roche, cat. no. 03117979001) **! CAUTION** Phenol is toxic when in contact with skin or if swallowed. Chloroform is harmful if inhaled or swallowed.
- 0.5-mm Glass beads (Biospec, cat. no. 11079105)
- 3 M NaOAc (pH 5.2–6.0)
- 70% and 100% EtOH
- *Escherichia coli* transformation competent cells (GC 10 cells; GeneChoice, cat. no. D-7L)
- Luria–Bertani (LB) plates containing 100 μg ml^{-1} ampicillin (see REAGENT SETUP)
- QIAprep Spin Miniprep Kit (Qiagen, cat. no. 27104)
- Sequencing primers (IDT) (see REAGENT SETUP) OM6189: CACAATATTTCAAGCTATACC; OM6373: CTCATCAACCAACGAAACGG
- T4 DNA ligase (with manufacturer's 10× ligation buffer; Roche, cat. no. 10481220001)
- Inverse PCR primers (IDT), see REAGENT SETUP and **Figure 1**: OM6609: CTTTTGGGTTATCACATTCAAC; OM6610: ATCGTAATTCACTACGTCAAC; OM6456: CCCATAACTGAATACGCATG; OM6458: AGGTTATGAGCCCTGAGAG
- REDTaq DNA polymerase (with manufacturer's buffers; Sigma, cat. no. D4309)
- QIAquick PCR Purification Kit (Qiagen, cat. no. 28104)
- BioPrime Array CGH Genomic Labeling Module (Invitrogen, cat. no. 18095-011)
- aCGH hyb buffer (Agilent, cat. no. 5188-5220)
- Yeast Whole Genome 4 × 44K ChIP-on-chip Microarray Kit (Agilent, cat. no. G4493A, design ID 014810)

EQUIPMENT

- Centrifuge (e.g., Eppendorf Centrifuge 5417R)
- Roller drum
- Bio-Rad *E. coli* pulser
- NanoDrop ND-1000 spectrophotometer
- Eppendorf BioPhotometer
- Thermal cycler (e.g., MJ Research PTC100)
- 30 and 37 °C incubator
- Liquid nitrogen and appropriate container
- Sonicator (e.g., Ultrasonic Processor XL2020)

REAGENT SETUP

Yeast lysis buffer 20 ml of 10% Triton (vol/vol), 10 ml of 10% SDS (wt/vol), 0.58 g NaCl, 1 ml of 1 M Tris (PH 8.0) and 200 μl of 0.5 M EDTA. Add water to 100 ml, filter sterilize.

LB ampicilin medium and plates Mix 10 g tryptone, 5 g yeast extract, 5 g NaCl (for plates, add 20 g agar), add water to 1 l, autoclave, cool and add 100 mg ampicillin.

YPD Mix 10 g yeast extract, 20 g peptone and 20 g glucose (for plates, add 20 g agar), dissolve in 1 l water, autoclave.

SC Mix 1.7 g yeast nitrogen base (Difco, cat. no. 233520), 5 g ammonium sulfate, 20 g glucose or Gal, various nutrient 'dropout' mixes (-His, -Ura, -Trp, -Ura -Trp; US Biologicals, use according to the manufacturer's instructions) (for plates, add 20 g agar), add water to 1 l, autoclave. For SC Glucose–His 5-FOA plates, after autoclave, add 1 g 5-FOA. Different SC plates are used to select for the markers on transformed plasmid. For example, SC Glucose–Ura plates are used to select yeast cells transformed with plasmid containing a *URA3* marker.

Primers Although we provide the sequences for inverse PCR and sequencing primers (see REAGENTS), oligos could also be designed by the user. All DNA primers should be synthesized at the 25-nmol scale. No purification other than standard desalting is necessary.

Primer design for cloning TF–Sir4 fusion construct Each primer (forward and reverse) is composed of two distinct sequences: the first (5′) 39 bp of each primer have the sequence of the regions flanking the Gal4DBD in the plasmid pBM4607, which are necessary to enable homologous recombination for cloning in yeast cells by gap repair[4,5]. The next ∼20 bp of each primer is a gene-specific sequence designed to amplify the open reading frame (ORF) so that it is intact and in frame with the Sir4-encoding sequences in pBM4607. We generally design primers to the first and

last 18–22 bp of the ORF. If the TF is to be fused to the N terminus of the Sir4 fragment (as we usually do), make sure to exclude the stop codon. To fuse the TF to the N terminus of Sir4 in plasmid pBM4607, the 39-bp homologous sequences are as follows (5′ to 3′): forward: ATACAATCAACTCCAAGCTT GAAGCAAGCCTCCTGAAAG reverse: TTTGGGTTTGCTAGAATTAGTAT CACTATGCGACACTCT.

PROCEDURE

Construct TF–*SIR4* fusion ● TIMING 5–7 d

1| Design primers to amplify the coding sequence of the TF of interest, as described in REAGENT SETUP.

2| Amplify the coding sequence of the TF with Phusion DNA polymerase (or other high-fidelity DNA polymerase) following the manufacturer's protocol. The PCR mix is made according to the following table, and should be prepared on ice.

Component	Amount (per reaction) (μl)	Final amount/concentration
5× Phusion HF buffer	5	0.5×
5× Phusion GC buffer	5	0.5×
10 mM dNTP mix	1	0.2 mM of each
Forward primer 25 μM	1	0.5 μM
Reverse primer 25 μM	1	0.5 μM
Yeast genomic DNA	1	10–100 ng
Phusion DNA polymerase	0.5	1 U
ddH_2O	35.5	
Total volume	50	

▲ **CRITICAL STEP** To avoid introducing mutations into TF coding sequence by PCR, always use high-fidelity DNA polymerase.

3| Program the thermocycler as follows:

Step	Temperature	Time	Cycles
1	98 °C	30 s	1
2	98 °C	10 s	
3	60 °C (variable depending on primer design)	30 s	
4	72 °C	15–30 s kb^{-1}	Go to Step 2 for 35 cycles
5	72 °C	5 min	1
6	4 °C	Indefinitely	1

4| Digest 1 μg of pBM4607 (contains the *GAL4DB–SIRr4* fusion with *TRP1* as the selectable marker) with 10 U of *Xho*I at 37 °C for 1 h.

5| Purify the linearized plasmid by gel electrophoresis. Run *Xho*I-digested pBM4607 on a 0.7% agarose gel (wt/vol) (containing 10 μg ml^{-1} ethidium bromide) at 130 V for 1 h. Cut out the DNA (should be in one band on the gel) and purify using the QIAquick Gel Extraction Kit (following the manufacturer's protocol).

6| Cotransform a *trp1* yeast strain with 10–30 ng of the linearized pBM4607 from Step 5 and all of the PCR product (usually >3 μg) from Step 3 as described in **Box 2**.

7| After 2-d incubation at 30 °C, multiple yeast colonies should be observed on SC Glucose–Trp plate. Pool eight Trp^+ colonies in 200 μl yeast lysis buffer. Extract DNA as described in **Box 3**. Resuspend DNA pellet in 100 μl ddH_2O.

8| Transform 1 μl extracted DNA into competent *E. coli* cells using Bio-Rad *E. coli* pulser following manufacturer's protocol and plate on LB + ampicillin plates.

9| Incubate plates at 37 °C overnight.

10| Pick four to eight *E. coli* colonies from the LB + ampicillin plate and culture each in 1 ml LB + ampicillin media. Incubate on a roller drum at 37 °C overnight.

11| Purify plasmid DNA from each culture using QIAprep Spin Miniprep Kit (follow the manufacturer's protocol) and determine the DNA sequence of the TF–*SIR4* junction using sequencing primers OM6189 and OM6373 (see REAGENTS).
■ **PAUSE POINT** Transform confirmed constructs in *E. coli* and store as glycerol stocks in −80 °C freezer, which can be kept for years.

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols

BOX 2 | YEAST TRANSFORMATION ● TIMING 2 d

The protocol described here is based on that provided in reference 24.
1. Start a 5 ml YPD culture of the working strain 1 d earlier. Incubate overnight on a roller drum at 200 r.p.m. and 30 °C.
2. The next day, pipette 100 μl cell suspension into 1 ml water in a spectrophotometer cuvette and measure OD at 600 nm using Eppendorf BioPhotometer. For most yeast strains, culture containing 1×10^6 cells per ml will give OD_{600} of 0.1.
3. Add 2.5×10^8 cells into 50 ml fresh YPD in a culture flask to give 5×10^6 cells per ml. Shake the culture at 30 °C and 200 r.p.m. for 3–5 h, until the cell density reaches $\sim 2 \times 10^7$ cells per ml.
▲ CRITICAL STEP Optimal cell density is critical to the transformation efficiency, do not use over-grown cells.
4. Spin down cells at 3,000*g* for 5 min and wash them with 10 ml sterile water. These cells are sufficient for ten transformations.
5. Aliquot cells for each transformation into 1.5-ml microcentrifuge tube. Spin down cells at 20,000*g* for 30 s and discard the supernatant. Make the total transformation mix first and then add 360 μl of the mix to each tube.

Component	Amount (per reaction) (μl)	Final amount/concentration
PEG 3500 50% (wt/vol)	240	33.3%
LiAc 1.0 M	36	0.1 M
Denatured (by boiling) SS-carrier DNA (10 mg ml^{-1})	10	27.8 ng μl^{-1}
DNA plus ddH_2O	74	
Total volume	360	

▲ CRITICAL STEP Be careful to pipette the correct volume of PEG, which is viscous.
6. Vortex the mixture vigorously and incubate the tube in a 42 °C water bath for 40 min.
7. Spin down the cells at 20,000*g* for 30 s and discard the supernatant. Add 100 μl ddH_2O into each tube and stir the pellet with pipette tip. Plate appropriate dilution of the cell suspension onto SC selection media. For example, if a plasmid containing *TRP1* marker was transformed, plate cells on SC Glucose–Trp media.
8. Incubate the plates at 30 °C for 2–3 d.

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols

Induction and selection of Ty5 transposition ● TIMING 10 d

12| Cotransform the plasmids containing the TF–*SIR4* fusion (with *TRP1* marker) and the Ty5 transposon (with *URA3* marker) into yeast strain YM7635 as described in **Box 2**. Use 0.1–0.5 μg of each plasmid for transformation. Remember to carry out a control experiment in parallel, as described in **Box 1**.
? TROUBLESHOOTING

13| Incubate for 2 d at 30 °C, after which time multiple yeast colonies containing both plasmids should be observed on SC Glucose–Ura–Trp plate.

14| From each plate, pick one colony and culture overnight in 5 ml SC Glucose–Ura–Trp media at 30 °C. Once the culture reaches an OD_{600} of 1 or higher, plate 500 μl of cells on each of 10 SC Glucose–Ura–Trp plates.

15| Grow at 30 °C for 1 d until a confluent lawn is formed.

16| Replica plate the cells onto SC Galactose–Ura–Trp plates to induce Ty5 transposition. Galactose will activate the *GAL1-10* promoter that drives the expression of Ty5. Keep plates at room temperature (22–25 °C) for 3 d.

17| Select for cells with Ty5 transpositions by replica plating onto SC Glucose–His plates. The integrated Ty5 transposon has a functional His marker, so only cells with transpositions will grow. Incubate plates for 2 d at 30 °C.
? TROUBLESHOOTING

BOX 3 | YEAST GENOMIC DNA EXTRACTION ● TIMING 1.5 h

1. Add 200 μl yeast lysis buffer, 200 μl phenol/chloroform/iso-amyl alcohol (25:24:1) and 200 μl of 0.5-mm glass beads to 50-μl cell pellet. Vortex for 5–10 min.
! CAUTION Phenol is toxic when in contact with skin or if swallowed. Chloroform is harmful if inhaled or swallowed.
2. Spin the tubes in a microcentrifuge at 20,000*g* for 10 min. Transfer the supernatant into a new 1.5-ml microcentrifuge tube.
▲ CRITICAL STEP Avoid the transfer of debris from the interface to reduce the contamination of protein in the extracted DNA.
3. Add 200 μl chloroform to the tube, vortex well and spin at 20,000*g* for 5 min. Transfer the supernatant into a new 1.5-ml microcentrifuge tube.
! CAUTION Chloroform is harmful if inhaled or swallowed.
4. Add 1/10 volume 3 M NaOAc (pH 5.2–6.0) and 2.5 volume of 100% EtOH. Vortex vigorously and keep at 80 °C for 30 min.
5. Spin the tube in a microcentrifuge at 20,000*g* for 10 min. A pellet of DNA should be visible.
6. Decant the ethanol and add 1 ml of 70% EtOH to the DNA pellet. Invert the tube several times, spin at 20,000*g* for 5 min.
7. Decant the 70% EtOH, vacuum dry the DNA pellet and resuspend DNA in 100 μl TE or ddH_2O. The DNA concentration should be $\sim$200 ng μl^{-1}.

18| Select colonies that have lost the Ty5-containing plasmid by replica plating onto SC Glucose–His 5-FOA plates. 5-FOA will counter-select the cells containing *URA3* gene. Incubate for 2 d at 30 °C.
? TROUBLESHOOTING

19| Harvest the cells from the SC Glucose–His 5-FOA plates with transposed Ty5 by adding 1 ml of YPD to each plate. Suspend the cells using a spreader and pipette the liquid into a 15-ml Falcon tube. Pool the cells from all ten plates into one 15-ml Falcon tube, which will yield ~8 ml cells in YPD.

20| Aliquot 50 µl cell pellet and extract genomic DNA as described in **Box 3**.
■ **PAUSE POINT** Freeze the remaining cells in liquid nitrogen and store in −80 °C freezer for as long as needed. Extracted DNA can be stored at −20 °C for several months before proceeding with next step.

Enzyme digestion, DNA fragment circularization and amplification ● TIMING 1.5 d

21| Digest 1 µg genomic DNA from Step 20 with *Taq*I, *Hin*P1I and *Hpa*II independently. For each digestion, add 2 µl 10× NEB buffer, 2 µl 10× BSA, 10 U restriction enzyme, 1 µg genomic DNA and add ddH_2O to total 20 µl. For *Hin*P1I and *Hpa*II digestions, incubate at 37 °C for 1 h. For *Taq*I digestion, incubate at 65 °C for 1 h.

22| Run 2 µl of each reaction on a 0.7% agarose gel (wt/vol) to confirm DNA digestion. A 200 bp to 5 kb smear should be observed.

23| Purify the DNA from each reaction using the QIAquick PCR Purification Kit, following the manufacturer's protocol. To elute the DNA from the column, apply 30 µl ddH_2O at the center of the column, let it sit on bench for 1 min and spin the column at 18,000*g* for 1 min. Measure the DNA concentration using the NanoDrop apparatus.

24| To circularize the digested fragments, prepare the following ligation reaction on ice and incubate at 15 °C overnight.

Component	Amount	Final
Digested DNA	50–100 ng	50–100 ng
10× T4 ligation buffer	10 µl	1×
T4 DNA ligase	1 Weiss U	1 U
ddH_2O	To 100 µl	

▲ **CRITICAL STEP** Do not use >100 ng digested DNA in ligation reaction, or intermolecular ligations will be favored over the desired intramolecular circularization.

25| For each digested and circularized sample, amplify the ligated products from 5 µl of the ligation reaction by PCR. Set up separate reactions with one pair of primers to amplify the genomic regions on the left side (primers OM6609 and OM6458, see REAGENTS) and with another pair of primers to amplify the right side (primers OM6610 and OM6456, see REAGENTS) of Ty5 (**Fig. 1**). Set up the reactions by mixing the following components on ice:

Component	Amount (per reaction) (µl)	Final amount/concentration
10× REDTaq buffer	5	1×
10 mM dNTP mix	1	0.2 mM of each
5 M Betaine	10	1 M
Forward primer 25 µM	1	0.5 µM
Reverse primer 25 µM	1	0.5 µM
Ligation mix from Step 24	5	2.5–5 ng
REDTaq DNA polymerase	2	2 U
ddH_2O	25	
Total volume	50 µl	

26| Program the thermocycler as follows:

Step	Temperature (°C)	Time	Cycles
1	93	2 min	1
2	93	30 s	
3	60	6 min	Go to Step 2 for 28–30 cycles
6	4	Indefinitely	1

27| Run 5 µl of the PCR products on a 0.7% agarose gel (wt/vol) and a 200 bp to 2 kb smear should be observed.

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols

28| Purify each PCR product with the QIAquick PCR Purification Kit (following the manufacturer's protocol) and measure the DNA concentration using a NanoDrop apparatus. For each TF, pool the same amount of DNA from each PCR product (total six PCR products from three different digested and self-ligated samples).

29| Prepare the control sample: shear 10 μg of yeast genomic DNA using sonicator. Using Ultrasonic Processor XL2020, shear DNA for 1 min at full power (level 10).
▲ CRITICAL STEP Keep the sample in ice-water bath during sonication; use a clamp to hold the tube in the ice-water bath so that the bottom of the tube sits 0.5–1.0 cm above the sonicator probe.

30| Run a portion of the sheared genomic DNA on a 0.7% agarose gel (wt/vol) to confirm DNA shearing. A 200 bp to 2 kb smear should be observed.
■ PAUSE POINT Sample and control DNA could be stored at −20 °C for several weeks before microarray hybridization.

Microarray hybridization and data analysis ● TIMING 3 d

31| Label both the PCR products (test DNA) and the sheared genomic (control) DNAs with Invitrogen's BioPrime Array CGH Genomic Labeling Module, using a different fluorophore (cy3 or cy5) for each. Follow manufacturer's protocol with the following exceptions/specifications: input mass for genomic DNA = 1.6 μg DNA/fluorophore/array; input mass for PCR products = 2.0 μg DNA/fluorophore/array.
▲ CRITICAL STEP Because this method can be adapted to different microarray platforms, the protocol for hybridization and data analysis may vary. Here, we provide a general overview of the protocol that we employ.

32 | Cohybridize labeled DNAs to Agilent yeast Whole Genome 4 × 44K microarrays in Agilent aCGH hyb buffer; characterize each experimental condition in triplicate, using three microarrays. Follow Agilent's aCGH hybridization protocol with the following exceptions/specifications: hybridization overnight (16–20 h) at 65 °C at oven rotation of 20 r.p.m.

33| Washing: B1. Wash 1 = 6× SSPE/0.005% N-lauroylsarcosine (wt/vol); B2. Wash 2 = 0.06× SSPE; B3. Used Agilent stabilization and drying solution (cat. no. 5185-5979). Scan the microarrays on Genepix 4000B Microarray scanner (Molecular Devices) to detect cy3 and cy5 fluorescence.

34| Analyze images using the Genepix v6.0 software package to obtain fluorescent intensities for each feature on the microarray. Use the ratio of the mean fluorescent intensities of the test over control channel to estimate the extent of enrichment of loci present in the test DNA, then rank the loci based on this mean ratio. Next, use the Gcn4–Sir4 positive control to select the appropriate intensity cutoff. We typically choose a cutoff that maximizes the true Gcn4 positives at a 2.5% false-positive rate. A list of true Gcn4 targets, as well as a list of genes that are not targeted by Gcn4, can be found in the supplementary material of Pokholok *et al.*[23] For a more detailed description of data analysis, please see Wang *et al.*[8] and accompanying supplemental information (http://www.genome.org/cgi/data/gr.6510207/DC1/1).
? TROUBLESHOOTING

? TROUBLESHOOTING
Troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting table.

Step	Problem	Possible reason	Solution
12	Few or no colonies after transformation	Cotransformation of two plasmids is inefficient	Use more competent cells and plasmid DNA for transformation
16–18	Bacterial contamination on plates	Plates can easily become contaminated during replica plating	Important to clean bench with ethanol before replica plating. Autoclave velvets wrapped in foil (no more than ten per package) thoroughly before use
17	No colonies growing on synthetic complete Glucose–His plates (selecting for cells with Ty5 transposition)	Homologous recombination between 5′ and 3′ Ty5 long terminal repeats in the plasmid with the calling card results in the deletion of Ty5	Before inducing transposition, verify that the strain carries an intact Ty5 calling card by a PCR assay using a pair of primers that amplify a region within Ty5
34	Results from the strain with transcription factor (TF)–Sir4 are similar to results from the control strain without TF–Sir4	This TF–Sir4 fusion construct is nonfunctional	Determine the sequence of the entire coding sequence of TF–Sir4 to ensure there are no significant mutations. Also, expression of the fusion protein can be confirmed by western blotting with anti-Myc antibody (the Myc tag is fused to Sir4 in pBM4607). If the TF–Sir4 coding sequence and protein expression are fine, perhaps the TF interferes with the function of the Ty5 integrase. Try fusing only the DNA-binding domain to Sir4

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols

TABLE 2 | Raw data from microarray hybridization

	Gene name	Systematic name	cy5/cy3 ratio
GAL4			
Strong target	GAL10_GAL1	chr2:278552-278611	2,113.88144
Strong target	GAL10_GAL1	chr2:278210-278269	1,913.56122
Strong target	GAL10_GAL1	chr2:278766-278825	739.809151
Intermediate target	GAL2	chr12:290045-290104	242.987142
Intermediate target	GAL2	chr12:289916-289975	150.780762
Intermediate target	GAL2	chr12:289271-289330	19.8380465
Nontarget	ACT1	chr6:54741-54800	0.0035058
Nontarget	ACT1	chr6:53476-53535	0.01081764
Nontarget	ACT1	chr6:54282-54341	0.01021794
GCN4			
Known target	CPA2_YMR1	chr10:632975-633034	146.171857
Known target	CPA2_YMR1	chr10:633354-633413	136.359583
Known target	CPA2_YMR1	chr10:633184-633243	75.1000869
Known target	PRM5_HIS5	chr9:142513-142572	37.9003228
Known target	PRM5_HIS5	chr9:142382-142441	37.6663545
Known target	PRM5_HIS5	chr9:142799-142858	18.3060131
Nontarget	ACT1	chr6:55296-55355	1.27141814
Nontarget	ACT1	chr6:55068-55127	0.69365667
Nontarget	ACT1	chr6:54741-54800	0.22661866

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols

ANTICIPATED RESULTS

To define a set of genomic regions that have a high probability of being adjacent to a 'calling card', we used the calling cards to identify targets of the well-characterized TF Gcn4 and empirically chose a cutoff that minimizes the rate of false negatives at a false-positive rate of 2.5% (ref. 8). (A list of genes known to be regulated by Gcn4 and a list of genes that are not regulated by Gcn4 was provided by Pokholok *et al.*[23]) For each experiment, we performed three technical replicates. Probes with fluorescence ratios above the cutoff in at least two of the three measurements were considered significant. We ignored data from probes that cover telomere regions because Ty5 can insert into these regions of the genome due to homologous recombination with Ty5 elements that reside there. We also excluded *HIS3* probes because *HIS3* sequences from the Ty5 calling cards are present in the inverse PCR product.

Gal4 and Gcn4 provide good positive controls for the method. Gal4–Sir4 leaves calling cards at *GAL1-10*, *GAL7*, *GAL3*, *GAL2*, *FUR4*, *GCY1* and *PCL10*, approximately in that order of abundance[3]. Since a large number of calling cards are deposited upstream of *GAL1-10* and *GAL7*, the probes for these regions are often saturated in the test channel on the microarray. Gcn4 has more targets than Gal4, and consequently Gcn4–Sir4 leaves calling cards at a larger number of places in the genome[3]. A list of real and false Gcn4 targets can be found at Pokholok *et al.*[23]. For both Gal4–Sir4 and Gcn4–Sir4, the false negatives should be ~49% at a false-positive frequency of 2.5%. The negative control strains (i.e., no TF–Sir4 fusion) will contain transpositions that localize largely to the telomeres, although we also observe some background transposition in regions of open chromatin. We generally observe very different patterns of transposition in the negative control than in samples with TF–Sir4 fusions. It appears that background transposition (e.g., to the telomeres) is largely suppressed when the Sir4 protein is tethered to a TF.

An example of the raw data from microarray hybridization experiments of Gal4 and Gcn4 is shown in **Table 2**. For a strong target (*GAL1-10*), an intermediate target (*GAL2*) and a nontarget (*ACT1*) of Gal4, two known targets (*CPA2* and *HIS5*) and a nontarget (*ACT1*) of Gcn4, the top three probes on the microarray of each promoter are listed. The exact cy5/cy3 ratio often varied between biological replicates, but the relative ranking of target genes in each experiment remained largely the same.

ACKNOWLEDGMENTS We are grateful to Dan Voytas (Iowa State University) for generously providing reagents and advice. This work was supported by the US National Institutes of Health grants R21RR023960 and 5P50HG003170-03 and funds provided by the James S. McDonnell Foundation.

Published online at http://www.natureprotocols.com/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Voytas, D.F. & Boeke, J.D. Ty1 and Ty5 of *Saccharomyces cerevisiae*. In *Mobile DNA II* (eds. Craig, N.L., Craigie, R., Gellert, M. & Lambowitz, A.M.) 631–662 (ASM Press, Washington, DC, 2002).
2. Xie, W. *et al.* Targeting of the yeast Ty5 retrotransposon to silent chromatin is mediated by interactions between integrase and Sir4p. *Mol. Cell. Biol.* **21**, 6606–6614 (2001).
3. Zhu, Y., Dai, J., Fuerst, P.G. & Voytas, D.F. Controlling integration specificity of a yeast retrotransposon. *Proc. Natl. Acad. Sci. USA* **100**, 5891–5895 (2003).

4. Ma, H., Kunes, S., Schatz, P.J. & Botstein, D. Plasmid construction by homologous recombination in yeast. *Gene* **58**, 201–216 (1987).
5. Wach, A., Brachat, A., Pohlmann, R. & Philippsen, P. New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast* **10**, 1793–1808 (1994).
6. Ren, B. *et al.* Genome-wide location and function of DNA binding proteins. *Science* **290**, 2306–2309 (2000).
7. Horak, C.E. & Snyder, M. ChIP-chip: a genomic approach for identifying transcription factor binding sites. *Methods Enzymol.* **350**, 469–483 (2002).
8. Wang, H., Johnston, M. & Mitra, R.D. Calling cards for DNA-binding proteins. *Genome Res.* **17**, 1202–1209 (2007).
9. Gao, X., Rowley, D.J., Gai, X. & Voytas, D.F. Ty5 gag mutations increase retrotransposition and suggest a role for hydrogen bonding in the function of the nucleocapsid zinc finger. *J. Virol.* **76**, 3240–3247 (2002).
10. Robertson, G. *et al.* Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* **4**, 651–657 (2007).
11. Johnson, D.S., Mortazavi, A., Myers, R.M. & Wold, B. Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* **316**, 1497–1502 (2007).
12. Lee, T.I. *et al.* Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804 (2002).
13. Harbison, C.T. *et al.* Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**, 99–104 (2004).
14. Boyer, L.A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
15. Zeitlinger, J. *et al.* Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev.* **21**, 385–390 (2007).
16. Schmid, M., Durussel, T. & Laemmli, U.K. ChIC and ChEC; genomic mapping of chromatin proteins. *Mol. Cell* **16**, 147–157 (2004).
17. Wilson, T.E., Fahrner, T.J., Johnston, M. & Milbrandt, J. Identification of the DNA binding site for NGFI-B by genetic selection in yeast. *Science* **252**, 1296–1300 (1991).
18. Deplancke, B., Dupuy, D., Vidal, M. & Walhout, A.J. A gateway-compatible yeast one-hybrid system. *Genome Res.* **14**, 2093–2101 (2004).
19. Deplancke, B. *et al.* A gene-centered *C. elegans* protein-DNA interaction network. *Cell* **125**, 1193–1205 (2006).
20. Vermeirssen, V. *et al.* Transcription factor modularity in a gene-centered *C. elegans* core neuronal protein-DNA interaction network. *Genome Res.* **17**, 1061–1071 (2007).
21. van Steensel, B. & Henikoff, S. Identification of *in vivo* DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat. Biotechnol.* **18**, 424–428 (2000).
22. van Steensel, B., Delrow, J. & Henikoff, S. Chromatin profiling using targeted DNA adenine methyltransferase. *Nat. Genet.* **27**, 304–308 (2001).
23. Pokholok, D.K. *et al.* Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517–527 (2005).
24. Gietz, R.D. & Woods, R.A. Yeast transformation by the LiAc/SS Carrier DNA/PEG method. *Methods Mol. Biol.* **313**, 107–120 (2006).

© 2008 Nature Publishing Group http://www.nature.com/natureprotocols