Genomic profiling of drug sensitivities via induced haploinsufficiency

Guri Giaever1, Daniel D. Shoemaker2, Ted W. Jones3, Hong Liang1, Elizabeth A. Winzeler1, Anna Astromoff1 & Ronald W. Davis1,3

Lowering the dosage of a single gene from two copies to one copy in diploid yeast results in a heterozygote that is sensitized to any drug that acts on the product of this gene. This haploinsufficient phenotype thereby identifies the gene product of the heterozygous locus as the likely drug target. We exploited this finding in a genomic approach to drug-target identification. Genome sequence information was used to generate molecularly tagged heterozygous yeast strains that were pooled, grown competitively in drug and analysed for drug sensitivity using high-density oligonucleotide arrays. Individual heterozygous strain analysis verified six known drug targets. Parallel analysis identified the known target and two hypersensitive loci in a mixed culture of 233 strains in the presence of the drug tunicamycin. Our discovery that both drug target and hypersensitive loci exhibit drug-induced haploinsufficiency may have important consequences in pharmacogenomics and variable drug toxicity observed in human populations.

Introduction

It is often assumed that one copy of a gene is adequate for the normal function of diploid organisms in all but a few cases. This assumption is based primarily on the fact that haploinsufficiency, in which loss of function of one gene copy leads to an abnormal phenotype, is rarely observed. In humans, for example, only a handful of genes have been identified as haploinsufficient1. These were identified because they resulted in developmental abnormalities or other severe disease. In a genome-wide analysis of Drosophila melanogaster, 56 haploinsufficient loci (approximately 0.4% of the Drosophila coding capacity) exhibited lethality or abnormality2. These examples demonstrate that, regardless of their frequency, haploinsufficient loci define a set of genes whose dosage and function are critical to the organism. In this study, we designed experiments to allow identification of haploinsufficient genes in yeast in a whole-genome manner. In addition to the identification of haploinsufficient genes under optimal growth conditions, we identified a class of genes that are haploinsufficient only under stressed conditions. This study focuses on the haploinsufficiency induced in the presence of a drug that targets the gene product of the heterozygous locus. In these experiments, the primary effect of drug is to decrease further gene function in heterozygous strains by inhibiting the remaining gene product, effectively widening the phenotypic window for identification of haploinsufficient genes.

The yeast Saccharomyces cerevisiae is a model organism for the study of haploinsufficiency due to its ease of genetic manipulation and the availability of genome sequence information, which allows for the systematic construction of heterozygous deletion strains in any essential or nonessential gene. Although comprehensive genome-wide surveys of gene dosage in yeast have not been performed, evidence that gene dosage can alter drug sensitivity has been demonstrated. Specifically, increased gene dosage (on the order of ten copies per cell) of a drug target has been shown to increase resistance of a yeast strain to the corresponding drug3,4. Here, the converse is demonstrated: decreased gene dosage of a drug target from two copies to one copy in heterozygous strains results in increased sensitivity, or drug-induced haploinsufficiency. Although the idea of increasing or decreasing copy number is not new, the study of heterozygous yeast strains with decreased copy number of a single gene has been largely unexplored due to the lack of discernible phenotypes in most heterozygous strains. We first performed a feasibility study to confirm increased sensitivity in individual strains, and then used this phenotype to screen heterozygous strains in parallel for drug sensitivity. The success of the parallel approach using 233 strains encourages efforts to scale to a genome-wide level.

Results

Individual heterozygous deletion strains exhibit drug-induced haploinsufficiencies

We constructed a set of heterozygous yeast strains carrying deletions in genes encoding known drug targets. The genes used in this study, which encode essential proteins involved in diverse cellular processes, include: HIS3, required for histidine biosynthesis; ALG7, required for glycosylation; RNR2, required for deoxyribonucleotide biosynthesis; TUB1 and TUB2, encoding tubulin subunits essential for cell division; and ERG11, crucial for ergosterol biosynthesis (Table 1). Each heterozygous strain was grown in the presence of sublethal concentrations of the drug that directly targets the protein encoded by the heterozygous locus. The growth rate of each heterozygote with respect to wild type provides a measure of the relative fitness, or sensitivity, of a strain in the presence of a drug. In these experiments, haploinsufficiency is manifested as a reduced fitness of a heterozygous strain compared with the wild-type strain under the same conditions.

1Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305-5307, USA. 2Rosetta Inpharmatics, 12040 115th Ave NE, Kirkland, Washington 98034, USA. 3Stanford DNA Sequencing and Technology Center, 855 California Ave, Palo Alto, California 94394, USA. Correspondence should be addressed to G.G. (e-mail: ggiaever@cmgm.stanford.edu).
As increased gene dosage of \( \text{ALG7} \) is known to confer increased resistance to tunicamycin \(^5,6\), we predicted that the decreased gene dosage in the \( \text{alg7} \)/\( \text{ALG7} \) heterozygous strain would confer increased sensitivity to tunicamycin. In the absence of tunicamycin, both \( \text{alg7} \)/\( \text{ALG7} \) and wild-type strains grew equally well, with a generation time of approximately 2 hours (Fig. 1a), whereas high concentrations of tunicamycin (2 \( \mu \)g/ml) prevented growth of both strains (Fig. 1c). At intermediate drug concentrations (0.5 \( \mu \)g/ml), however, we saw a difference in growth rate between the \( \text{alg7} \)/\( \text{ALG7} \) strain and wild type (Fig. 1b), demonstrating that the \( \text{alg7} \)/\( \text{ALG7} \) strain was more sensitive to tunicamycin than wild type.

We also tested the effects of the microtubule-depolymerizing agent benomyl on a \( \text{tub1} \)/\( \text{TUB1} \) heterozygote strain. In the absence of benomyl, the \( \text{tub1} \)/\( \text{TUB1} \) strain grew slower than the wild-type strain (Fig. 2a), as previously reported\(^7\). Hence, the level of Tub1p in the heterozygote is not sufficient for wild-type growth, and this strain is haploinsufficient for growth in the absence of benomyl. This inherent haploinsufficiency does not affect our analysis, however, as at intermediate concentrations of drug (in this example 25 \( \mu \)g/ml benomyl) growth of wild type is minimally affected, whereas growth of the \( \text{tub1} \)/\( \text{TUB1} \) strain is inhibited (Fig. 2b).

We carried out similar growth rate analyses for several other heterozygous strains using known inhibitors (Table 1). Each experiment was tested using at least two individual transformants, which all showed induced haploinsufficiency in response to the appropriate drug. Furthermore, when the \( \text{his3} \)/\( \text{HIS3} \) heterozygote strain was transformed with a single copy of plasmid-borne \( \text{HIS3} \), it became more resistant to the inhibitor 3-amino-triazole, confirming that the affected locus is responsible for the observed phenotype (data not shown). These results demonstrate that increasing or decreasing gene dosage by a single gene copy results in detectable phenotypes.

### Parallel identification of drug targets

The results from the individual strain study suggest that if a heterozygous strain exhibits increased sensitivity to a drug, then the heterozygous gene may encode the drug target. We incorporated this prediction into a parallel screen for the identification of drug-sensitive strains. Mixed cultures of heterozygous deletion strains were grown in the presence of drug. Strains that grew more slowly with respect to wild type represented those most sensitive to drug and indicated the possible identification of a drug target. We tested the feasibility of this parallel approach using strains tagged with ‘molecular bar codes’ that allow quantitative measurement of individual fitness when strains are grown competitively\(^8\).

We analysed a pool of 12 heterozygous strains in the presence of the drug tunicamycin, a well-characterized glycosylation inhibitor\(^9\). Each strain was uniquely tagged using a 20-base oligonucleotide bar code, as previously described\(^9\) (Fig. 3). Equal numbers of cells from each of 12 heterozygous strains were pooled and grown in rich medium in the presence or absence of 0.5 \( \mu \)g/ml tunicamycin (the concentration at which there was the greatest fitness differential between the individual \( \text{alg7} \)/\( \text{ALG7} \) strain and wild type). Aliquots of cells were taken from the pool over time and genomic DNA was isolated and used as template for PCR amplification of the oligonucleotide tags by two common primers, one incorporating a 5’ fluorescent label in the process. Subsequent hybridization of the DNA tags to a oligonucleotide array, followed by laser scanning, allowed quantitative analysis of the relative abundance of each strain in the pool. Each position on the oligonucleotide array contains a complementary oligonucleotide corresponding to a unique heterozygote strain, and the fluorescence intensity at this position reflects the relative abundance of each individual heterozygous strain.

At time zero in the 12-strain experiment, each strain was present in equal amounts and the fluorescent signals on the oligonucleotide array were of approximately equal intensity (Fig. 4).

**Table 1 • Genes encoding known drug targets used in individual strain analysis**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein; protein function</th>
<th>Inhibited by</th>
<th>Haploinsufficiency(^\text{a})</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIS3</td>
<td>imidazolesuglycerol-phosphate dehydratase; required for histidine biosynthesis</td>
<td>3-amino-triazole(^{21})</td>
<td>100 mM</td>
</tr>
<tr>
<td>ALG7</td>
<td>Asn-linked glycosyl transferase; required for Asn-linked glycosylation</td>
<td>tunicamycin(^9)</td>
<td>0.6 ( \mu )M</td>
</tr>
<tr>
<td>RNR2</td>
<td>small subunit of ribonucleotide reductase; required for nucleotide biosynthesis</td>
<td>hydroxyurea(^{22})</td>
<td>50 mM</td>
</tr>
<tr>
<td>TUB1</td>
<td>( \alpha )-tubulin structural protein; essential component of mitotic apparatus</td>
<td>benomyl(^7)</td>
<td>85 ( \mu )M</td>
</tr>
<tr>
<td>TUB2</td>
<td>( \beta )-tubulin structural protein; essential component of mitotic apparatus</td>
<td>benomyl(^7)</td>
<td>85 ( \mu )M</td>
</tr>
<tr>
<td>ERG11</td>
<td>cytochrome P450 lanosterol 14a-demethylase; required for biosynthesis of ergosterol</td>
<td>fluconazole(^{23})</td>
<td>40 ( \mu )M</td>
</tr>
</tbody>
</table>

\(^\text{a}\)Haploinsufficiency: concentration of drug where O.D.\(_{600}\) wild type/O.D.\(_{600}\) heterozygote was greatest. Gene information obtained from the Saccharomyces Genome Database (http://genome-www.stanford.edu/Saccharomyces).

---

**Fig. 1** Tunicamycin sensitivity of \( \text{alg7} \)/\( \text{ALG7} \). Growth of \( \text{alg7} \)/\( \text{ALG7} \) and \( \text{alg7} \)/\( \text{ALG7} \) wild-type strains (O.D.\(_{400}\)) as a function of time. \( \text{a} \), 0 \( \mu \)g/ml tunicamycin; \( \text{b} \), 0.5 \( \mu \)g/ml tunicamycin; \( \text{c} \), 2.0 \( \mu \)g/ml tunicamycin. Drug-induced haploinsufficiency is seen in (b), where an intermediate concentration of 0.5 \( \mu \)g/ml tunicamycin reveals a difference in drug response between the two strains.
the 0.5-µg/ml tunicamycin culture, the hybridization signal on the oligonucleotide array representing the alg7/ALG7 strain was diminished at 22 hours, and undetectable by 48 hours. All other strains in the pool remained at equal intensity (and therefore equal abundance) over the course of the experiment, indicating that they were relatively unaffected by the drug. These results demonstrated that a known drug target could be identified in a pool of 12 heterozygous strains.

The 12-strain study was expanded using a set of 233 unique heterozygous yeast strains constructed by a consortium of laboratories (E.A.W. et al., manuscript in preparation; http://sequence-www.stanford.edu/group/yeast_deletion_project/deletions3.html) and was performed in the same manner as the 12-strain pool experiment. The analysed data from this study are shown (Fig. 5). Each bar in the graph reflects the sensitivity of a heterozygous strain grown with and without 0.5 µg/ml tunicamycin. The greater the positive value of the bar, the greater the rate the strain is diminishing from the pool. The greater the positive value of the bar, the greater the rate the strain is increasing in the pool. The unlabelled positive and negative bars (Fig. 5) define the background fluctuation in a given experiment due to natural variability and two 50-bp sequence at both ends homologous to upstream and downstream sequences that flank the ORF of the target gene, allowing for site-specific mitotic recombination20. DNA was extracted and all tags were amplified in a single PCR reaction using common primers. These PCR products were then hybridized to an oligonucleotide array carrying the complement of each tag to determine the relative abundance of each strain in the pool (Fig. 4).

Neither strain exhibited increased sensitivity to the unrelated drugs hygromycin B and fluconazole. A conventional halo assay revealed extreme sensitivity to tunicamycin for ymr007w/YMR007W, but relatively little increased sensitivity for the alg7/ALG7 and ymr266w/YMR266W strains (Fig. 6). The discrepancy between the relative sensitivities of the strains in liquid culture versus growth on plates probably reflects the increased sensitivity of the direct competitive growth in combination with the differences in the physiological state of the cells. It should be noted that the amount of drug used for the plate assay was 40 times that used for the liquid culture assay.

**Discussion**

We tested six heterozygous strains carrying deletions in known drug targets for induced haploinsufficiency, all of which revealed induced haploinsufficiency in the presence of a drug that targets the gene product of the heterozygous locus. In each case, the result was highly specific, as we saw no haploinsufficiency when these strains were tested with other drugs. These strains define a class of genes that exhibit induced haploinsufficiency in the presence of a drug. Induced haploinsufficiency may also be revealed under other general stress conditions, as well as in conditions that specifically stress a particular gene product. This suggests that a variety of environmental perturbations, when combined with genetic lesions, will reveal previously undetected haploinsufficiencies.

Drug-sensitivity profiling of 233 heterozygous strains in the presence of tunicamycin identified three drug-sensitive loci: (i) the ALG7 locus, encoding the known target of tunicamycin; (ii) YMR007W, encoding a 126-aa protein of unknown function; and (iii) YMR266W, encoding a 953-aa protein with homology to the multi-facilitator superfamily11 (MFS). Because both the ymr007w/ymr007w and ymr266w/ymr266w strains were also...
drug-sensitive, these loci were ruled out as possible drug targets because the proteins they encode are absent in these homozygotes. As YMR266W shares homology with genes encoding the MFS family, it may encode a permease or another protein involved in net drug export. Further genetic analysis will allow discrimination of genes conferring sensitivity into categories that are either directly connected to the drug target(s), or involved in drug availability. Those genes directly connected to the drug target will aid in elucidation of the drug target pathway. The unambiguous identification of the known drug target in these experiments encourages future efforts to scale to a genomic level, which would permit screening against all 6,000+ potential drug targets simultaneously.

In addition to identifying drug-sensitive loci, all strains that are haploinsufficient in the absence of drug will be identified. We identified one such strain in the 233-strain pool experiment. This strain carries a deletion in YMR242C, which has been classified by sequence homology as a gene that codes for a cytoplasmic ribosomal protein. In the presence of tunicamycin, this strain did not exhibit a slow-growth phenotype. Because the average generation time for most of the strains in the presence of tunicamycin was approximately twice greater than in the absence of drug, the decreased Ymr242cp protein level may no longer be inhibitory for growth under these conditions. In general, strains haploinsufficient in the absence of drug identify gene products (including those that may be essential under defined growth conditions) that may be good candidates for drug targets, because in these strains small changes in protein levels result in a substantial decrease in fitness.

Currently, drug discovery is largely driven by combinatorial chemistry followed by high throughput in vitro screening of compounds against a preselected target. This method, though effective, requires that molecular targets be chosen a priori. An advantage of our approach to drug-target identification is that no prior knowledge of the target is required, and only those targets that affect the fitness of the organism will be identified. Furthermore, because this is an in vivo assay, only drugs that enter the cell and are not significantly metabolized are identified. Finally, novel anti-fungal targets, as well as those targets in critical common pathways that yeast share with higher eukaryotic cells, will be revealed.

This approach will likely identify all targets of any particular drug, toxin, or complex mixture (such as a natural product extract) that result in haploinsufficiency, as all potential targets are individually evaluated for increased sensitivity. This is an advantage over a complementary in vivo approach.
whereby targets are identified by their ability to confer resistance to a strain when present at high copy 3,4. In this technique, a more sensitive target present at wild-type levels may mask any resistance conferred by overexpression of an additional target. The less-sensitive target in this case would therefore escape detection.

This approach may also be applied to other organisms. It is not necessary to specifically delete a gene and replace it with a molecular bar code as described in these experiments. It should be possible, in principle, to introduce molecularly bar-coded transposons that would randomly mutagenize genes by insertion throughout the genome, although it is difficult to obtain strains with only single hits following transposition. Location of the insertion may then be accomplished using transposon-specific primers 12,13. Experimentally, there is also no requirement that the transposition event be single hits following transposition. Location of the transposons that would randomly mutagenize genes by insertion is possible, in principle, to introduce molecularly bar-coded transposons that would randomly mutagenize genes by insertion throughout the genome, although it is difficult to obtain strains with only single hits following transposition.

Fig. 6 Halo assays of tunicamycin-sensitive strains. An overnight culture (200 µl) was spread onto YPD plates. Tunicamycin (10 mg/ml) was spotted onto filter paper (~1/4” in diameter) that was placed on a lawn of the four heterozygous strains indicated. Plates were incubated at 30 °C for 3 d.

Methods

Regents. Unless otherwise specified, all reagents were from Sigma. Fluconazole was the gift of J. DeRisi.

Yeast strains. The yeast strain S288c, a wild-type isolate, was used as the parent for mutant construction in the individual growth study as well as in the 12-strain pool experiment. A derivative of S288c, BY4743, was used as the parent in the 233-strain pool experiment (http://sequence-ww.stanford.edu/group/yeast_deletion_project/deletions3.html). Culture media and conditions were as described 19.

Individual growth rate analysis. Growth rates of heterozygous deletion strains were measured and compared with the growth rate of wild type at varying drug concentrations. Cells were diluted from an overnight culture to an O.D. 600 of ~0.01 (~2.10^5 cells/ml) and allowed to grow until the O.D. 600 reached ~0.05 (~10^6 cells/ml; t=0), ensuring that the cells were in logarithmic phase. Drug was then added and growth rate was measured as the optical density of cells (O.D. 600) as a function of time (h) in rich medium 19 (YPD).

Pooled growth studies. Equal numbers of cells (~5x10^5) of each heterozygote strain were pooled and grown in a culture (100 ml) in the presence and absence of drug (starting O.D. 600=0.05). The doubling time in the absence of tunicamycin was ~1.5 h, and in the presence of tunicamycin ~3.0 h. Aliquots of ~2x10^7 (~1 O.D. 600) of cells were sampled over time. When the pooled culture reached ~3x10^5 cells (or an O.D. 600 of ~1.3 in 100 ml) the cultures were diluted back to 0.05 O.D. 600 (representing ~5x10^5 cells of each strain) as needed to maintain logarithmic growth. Because the time course of these experiments is within 15 generations of the pooled culture, the possibility of second-site suppressor mutations overtaking the culture is minimal. Both the 12-strain pool and the 233-strain pool were performed twice, each time yielding similar results. In addition, any strain appearing deficient in these experiments was verified by growth in individual cultures.

Construction of heterozygous deletion strains. Construction of strains used in the 12-strain pool was as described 18. An explanation of the construction of deletion strains used in the 233-strain study is available (http://sequence-ww.stanford.edu/group/yeast_deletion_project/deletions3.html). A note of caution regarding strain genotypes: it is possible, although rare, that some small fraction of these deletion strains carry second-site mutations resulting from transformation.

Hybridization. Analysis was as described 18.

Data analysis. Analysis of oligonucleotide array fluorescent signal data consisted of two steps: (i) normalization of data to achieve equality of background and maximal signals on each array; and (ii) analysis of the decrease (or increase) of signal intensity for each tag sequence over time. Equalization of signal strength relied on a consensus score based on the fact that for most oligonucleotide array sites, the signal strength did not vary between time points. The slope of a linear regression of the log of the above-background signals from each individual molecular bar code as a function of time identified strains that were either increasing or decreasing in abundance over time.

Acknowledgements

We thank Affymetrix for their help with all aspects of the oligonucleotide array technology used here. Specifically, M. Mittmann and D. Morris for oligonucleotide selection and chip design; E. Gentilen for array synthesis; and D. Lockhart for helpful discussions. We thank C. Nislow and I. Herskowitz for critical reading and helpful comments on the manuscript and all members of the S. cerevisiae Deletion Consortium for their efforts. Supported by a grant from Rhone-Poulenc Rorer and NIH grant H600198.

Received 1 December 1998; accepted 2 February 1999.