Yeast Augmented Network Analysis (YANA): a new systems approach to identify therapeutic targets for human genetic diseases [version 1; referees: 2 approved]

David J. Wiley¹, Ilona Juan¹, Hao Le², Xiaodong Cai³, Lisa Baumbach⁴, Christine Beattie², Gennaro D’Urso¹

¹Department of Molecular and Cellular Pharmacology, University of Miami Miller School of Medicine, Miami, FL, 33136, USA
²Department of Neuroscience, Ohio State University, Columbus, OH, 43210, USA
³Department of Engineering, University of Miami, Miami, FL, 33124, USA
⁴Integrated Functional Cancer Genomics, TGEN, Phoenix, AZ, 85004, USA

Abstract
Genetic interaction networks that underlie most human diseases are highly complex and poorly defined. Better-defined networks will allow identification of a greater number of therapeutic targets. Here we introduce our Yeast Augmented Network Analysis (YANA) approach and test it with the X-linked spinal muscular atrophy (SMA) disease gene UBA1. First, we express UBA1 and a mutant variant in fission yeast and use high-throughput methods to identify yeast genetic modifiers of UBA1. Second, we analyze available protein-protein interaction network databases in both fission yeast and human to construct UBA1 genetic networks. Third, from these networks we identified potential therapeutic targets for SMA. Finally, we validate one of these targets in a vertebrate (zebrafish) SMA model. This study demonstrates the power of combining synthetic and chemical genetics with a simple model system to identify human disease gene networks that can be exploited for treating human diseases.

Corresponding author: Gennaro D’Urso (gdurso@iomics.org)

How to cite this article: Wiley DJ, Juan I, Le H et al. Yeast Augmented Network Analysis (YANA): a new systems approach to identify therapeutic targets for human genetic diseases [version 1; referees: 2 approved] F1000Research 2014, 3:121 (doi: 10.12688/f1000research.4188.1)

Copyright: © 2014 Wiley DJ et al. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

Grant information: This work was supported by a grant from the Muscular Dystrophy Association (MDA186435). (Christine Beattie grant support RO1NS050414 (C.E.B.) with additional support from P30NS045758). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: No competing interests were disclosed.

Introduction

Many disease-associated genes have been identified, yet most genetic diseases remain untreatable. One path to treatment is to develop extensive genetic networks in which human disease genes function (or dysfunction) and then target therapies to the genes identified in those networks. Eventually, an individual’s own genotype for proteins in the network may also be considered in the therapeutic options. Developing gene networks is already widely recognized as a powerful approach to identify new drug targets. Gene networks are based on the principle that networks contain proteins that interact (physically and/or functionally) and that these interactions govern most, if not all, cellular functions. Importantly, gene network interactions are often conserved in different organisms even though the output of the networks may differ. Candidate networks can be generated from studies in model organisms and then extrapolated to human cells. One can expect that, for disease gene networks, some of the interacting genes might modulate the disease phenotype thus representing potential therapeutic targets.

The direct identification of disease networks from studies in human cells is, of course, challenging because of the large number of potential proteins and incomplete knowledge of how the activity of any one protein may affect the network output. Analysis of networks would benefit from studies in a tractable model organism amenable to high throughput methods. Given that network interactions may be conserved between humans and model organisms, even while the outputs may differ, we hypothesized that we can identify human disease networks in the fission yeast Schizosaccharomyces pombe, a simple, genetically tractable model organism, coupled to existing protein-protein interaction databases. We can then transfer that network knowledge to human cells with the goal of identifying therapeutic target genes.

How valid is the approach of using fission yeast networks to identify human disease gene networks? First, the proteins that control most core cellular functions in fact are evolutionarily conserved, underscoring the “deep homology” that exists between all living organisms. For example, in S. pombe >65% of the genome is orthologous to human (http://orthoml.org/orthoml/ and http://www.pombase.org). Second, a number of the cellular processes implicated in human disease, e.g. vascular transport, protein folding, metabolism, and RNA processing, are also evolutionarily conserved and highly interconnected. Importantly, the value of a simple model organism for discovering “druggable” genetic pathways has recently been demonstrated for the budding yeast Saccharomyces cerevisiae, and thus extending the study of human disease genes to fission yeast seems promising.

Our approach (Figure 1) involves the expression of human disease associated genes in S. pombe and then the analysis of their effect on yeast fitness (growth). We performed high-throughput synthetic genetic array (SGA) screens to identify the fission yeast genetic modifiers that alter this effect, as measured by a simple yeast growth assay. The genetic modifiers are then assembled into human disease gene networks or clusters (using protein-protein interaction datasets from both S. pombe and humans). Any modifiers or genes in the networks or clusters represent potential therapeutic targets. This unbiased high-throughput approach could be widely and rapidly applied to many different disease-associated genes at relatively low cost.

To demonstrate the power of YANA, we considered spinal muscular atrophy, a common neurodegenerative disease affecting approximately 1/6000 births worldwide and the number one genetic cause of infantile death in the United States. Most cases (>90%) are caused by deletion of SMN1, a gene encoding the survival motor neuron (SMN), a protein involved in the assembly of spliceosomal small nuclear ribonucleoproteins (snRNPs). In contrast, X-linked spinal muscular atrophy is caused by mutations in UBA1, a gene encoding the ubiquitin activating (UBA) enzyme 1, an E1 ubiquitin ligase. Consistent with a role of ubiquitination in spinal muscular atrophy and SMN biology, it was recently shown that ubiquitin-mediated proteolysis regulates SMN stability. Therefore, we systematically investigated the genetic network for UBA1 to better understand the potential connections between UBA1, its modifiers and the spinal muscular atrophy phenotype. Using YANA we identified several potential therapeutic targets and validated one of these targets in a SMA vertebrate (zebrafish) model.

Methods

Query Strain Creation

UBA1 was cloned into a pENTR/D-TOPO vector (Life Technologies, Cat # K2400-20) from cDNA (Origene, Cat. # SC320329) following the manufacturer’s protocols. The following primers were used: forward primer: 5′-CACATGTCAGCTGCGCCGCGG-3′; reverse primer: 5′-TCAGCGGATGTTGATCCGACG-3′. Genetic insertion was confirmed by sequencing (http://sylvester.org/shared/resources/oncogenomics). UBA1 (G1617T), mutUBA1, was created by site directed mutagenesis using a QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies, Cat. # 210518; the detailed protocol is available in the kit). The primers used were: forward primer: 5′-GCAGCTTGCCCAAATATCTACACATATCCGGG-3′; reverse primer: 5′-CCGGATATGTGGATTAATTTG-3′.

Gateway reactions (Gateway Cloning Protocols: http://www.lifetechnologies.com/us/en/home/life-sciences/cloning/gateway-cloning/protocols.html#lr) were then performed to flip the UBA1 genes into destination vectors, to create N-terminal HA-tagged UBA1 under the control of the nmt1 promoter (LR Clonase II from Life Technologies, Cat # 11791020). The newly generated expression vectors were then integrated into an h leu1-32 ura4-D18 Ade6-M210 S. pombe strain (PN572) to create a wtUBA1 query strain (h integrated pjk148-nmt1-HA-UBA1-nmt1′′′-leu1-32 ura4-D18 Ade6-M210) and a mutUBA1 query strain (h integrated pjk148-nmt1-HA-UBA1(G1617T)-nmt1′′′-leu1-32 ura4-D18 Ade6-M210). All media, growth conditions, and genetic manipulations were as previously described.

Western Blot and Growth Curves

Strains containing wtUBA1 and mutUBA1 were grown exponentially in PMG media (Sunrise Scientific Cat. #2060, keeping the OD 0.4 below 0.4) or eight generations at 32°C and then induced for expression of the HA-tagged wtUBA1 and mutUBA1 by washing the cells three times with sterile water to remove thiamine. Cells were then grown exponentially for 16 hours and then lysed using a FastPrep 120 bead beater (MP Biomedical), followed by boiling in sample buffer (2x Laemmli Sample Buffer, Bio-Rad Cat. #161-0737) and then clarified by centrifugation. The expression of UBA1 was confirmed by Western Blot analysis using an anti-HA antibody at 1:2000 dilution (Covance, Cat #MMS-101P, AB_10063488) and standard procedures. Growth curve analysis was completed in the
presence and absence of thiamine (final concentration, 15 μM; Sigma, Cat. #T4625). Removal of thiamine was achieved by washing the cells three times with sterile water. Cells were incubated in the absence of thiamine for 22 hours under exponential growth, diluted and 120 μl added to the Tecan Infinity F200 plate readers (starting OD of 0.05) for growth curve analysis with an n=5 for each sample. For validation experiments, *cul3* and *gsk3* were not pre-induced, while *pub1* was pre-induced for 22 hours. The deletions of *cul3*, *pub1*, and *gsk3* were confirmed by PCR analysis of genomic DNA from each strain, as previously described.

**SGA screening**

Query strains were grown in liquid media and then pinned to agar in a 384-format using a RoToR HDA (Singer Instruments). The query was then crossed to the *S. pombe* haploid deletion library (Bioneer, Version 3.0 equivalent) on SPAS media (details can be found at http://www-bcf.usc.edu/~forsburg/media.html) using a modified SGA procedure. For germination, four replicates of each cross were pinned to a 1536 format on selective PMG media containing thiamine, adenine (225 mg/L, Sigma Cat.#A8751), leucine (225 mg/L, Sigma Cat. #L8912), and the antibiotic G418 (150 μM, Gold Bio Cat. #G-1418) (PAUT+G418). After three days, the plates were then pinned to both PAUT+G418 (non-inducing) and PAU+G418 (inducing) plates. The colonies growing on the PAUT+G418 plates were documented on a flatbed scanner for the next three consecutive days. The PAU+G418 plates were grown for two days, and then re-pinned to fresh PAU+G418 plates and documented over the next three consecutive days. Based on the growth characteristics of the wtUBA1 strain, the plates were then pinned to fresh PAU+G418 plates and documented for an additional three days.

**Hit analysis**

The documented plates were analyzed for ‘Hits’ representing a growth defect (SL) or growth suppressor (SS) using ScreenMill software. The software is used to quantify colony size for each individual cross and then to normalize the quantified plates with and without replicate exclusion for each quadruplicate of the query crossed to a specific deletion strain. Data was then compared (between non-induced versus induced) and ranked in Excel (P≤0.05) (Microsoft). The orthologs were then identified based on curated data from PomBase (www.pombase.org; build 2013-11-11-v1), OrthoMCL (http://orthomcl.org/orthomcl/; Version 5), InParanoid8
Bio-Informatics

The *S. pombe* and human ‘Hits’ were analyzed in String (http://www.string-db.org) to map protein-protein interactions limited to data from experiments at the highest confidence level (0.900), and named *pombe* primary and human primary. The *S. pombe* primary was then extended by adding first degree neighbors to the original *pombe* ‘Hits’, keeping all interactions based on data from experiments at the highest confidence level (0.900). From the extended *pombe* network interaction map, new ‘Hits’ were extracted that corresponded to at least two previous ‘orphans’ (‘Hits’ that were not previously mapped) that interact through a nearest neighbor. All ‘Hits’ were then analyzed through PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) for relevance in terms of SMA, with special focus on the ‘Hits’ in the interaction maps.

Morpholino (MO) injection and drug test

Transgenic Tg(mnx1:0.6hsp70:GFP)os26 embryos that express GFP in ventrally projecting motor axons, referred to as Tg(mnx1:GFP) embryos, were used for all zebrafish experiments. Embryos were staged according to Fritz et al.23 All fish were grown and maintained in the Ohio State University (OSU) zebrafish facility following established protocols and OSU animal welfare guidelines as stated in Dr. Beattie’s animal protocol (On file at OSU: 2009A0141-R1). Tg(mnx1:GFP) embryos were injected with 4 ng *smn* MO at the one- to 2-cell stage to knock down Smn as previously described24. At 10 hours post-fertilization (hpf) injected embryos (in their chorions) were placed in Petri dishes in fish water (60 µg/ml Instant Ocean® sea salts) containing compounds in 0.75% dimethyl sulfoxide (DMSO) or 0.75% DMSO only and incubated at 28.5°C in incubator until 28 hpf. Tg(mnx1:GFP) embryos at 28 hpf were anesthetized with tricaine (250 µg/ml, Sigma A-5040) and fixed overnight at 4°C in 0.75% dimethyl sulfoxide (DMSO) or 0.75% DMSO only and incubated at 28.5°C in incubator until 28 hpf.

Zebrafish western blot

Samples of zebrafish embryos injected with 4 ng *smn* MO at the one- to two-cell stage and uninjected embryos were treated from 10 to 28 hpf with DMSO or with DMSO and compounds as described above were collected for western blot.

Samples were generated by boiling 25 identically treated embryos in 75 µl blending buffer (63 mM Tris (pH 6.8), 5 mM EDTA, 10% SDS), 10 µl (equivalent to 3 embryos or 75 µg of protein) were added to 10 µl of sample buffer (100 mM Tris (pH 6.8), 0.2% bromophenol blue, 20% glycerol, 200 mM dithiothreitol) and run on a 10% polyacrylamide gel, blotted to nitrocellulose, probed with mouse anti-Smn (1/500; MANSMA12, a gift from Dr. G.E. Morris or anti HuD (1/1000; Santa Cruz Cat#sc-28299, AB_627765) and detected by chemiluminescence of bound HRP-conjugated mouse antibody. Blots were stripped and re-probed with mouse anti-β-actin (1/1000; Santa Cruz Cat #sc-47778, AB_626632).

Results

We expressed the human wildtype *UBA1* (wt*UBA1*) and the disease-causing variant 1617*UBA1* (mut*UBA1*) in *S. pombe* under the control of a regulatable nmt1 promoter27. We identified strains with stable integrated human *UBA1* genes expressing equivalent levels of *UBA1* proteins (Figure 2A). Under non-inducing conditions (+ thiamine) all *UBA1*-containing strains had a similar growth rate to the control non-UBA1-containing wildtype strain (Figure 2B). We
then found that cells expressing human wtUBA1 (fission yeast cells grown in media lacking thiamine to induce UBA1 expression), experienced a decrease in cell growth. No growth defect was observed for yeast cells expressing mutUBA1 when compared to the control (Figure 2B).

We set out to identify yeast gene mutations that either enhance or suppress cell growth dependent upon expression of either wt or mutUBA1 (i.e., epistatic modifiers). We used automated genetics to introduce the wt and mutUBA1 genes into a deletion strain collection corresponding to more than 90% of the non-essential genes in S. pombe (ca. 80% of the complete genome). We then tested each of the >7000 unique strains expressing either wt or mutUBA1 combined with a specific gene deletion, for growth properties. Each strain was scored for significance of growth difference compared to a control, and the top hits (P<0.05) were assembled into a database (Table 1). The categories of growth phenotypes were those that were synthetic lethal for wildtype or mutant (wt-SL, or mut-SL) or synthetically suppressed for wildtype or mutant (wt-SS, or mut-SS). We identified 173 UBA1 or mutUBA1 modifiers. Notably,

### Table 1. S. pombe SGA top hits.

The top hits for wtUBA1 SL, wtUBA1 SS, mutUBA1 SL, and mutUBA1 SS were identified by bioinformatic analysis as described in the materials and methods section.

<table>
<thead>
<tr>
<th></th>
<th>wtUBA1 SL</th>
<th>wtUBA1 SS</th>
<th>mutUBA1 SL</th>
<th>mutUBA1 SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SPAC24H6.03</td>
<td>SPBC365.06</td>
<td>SPBC8D2.01</td>
<td>SPAC11G7.02</td>
</tr>
<tr>
<td>2</td>
<td>SPAC1687.15</td>
<td>SPAC11G7.02</td>
<td>SPBC6B1.10</td>
<td>SPCC13B11.01</td>
</tr>
<tr>
<td>3</td>
<td>SPAC3G9.07c</td>
<td>SPCC31H12.05c</td>
<td>SPBC216.05</td>
<td>SPBC2G2.03c</td>
</tr>
<tr>
<td>4</td>
<td>SPBC6B1.10</td>
<td>SPBC31F10.13c</td>
<td>SPCC1795.01c</td>
<td>SPBC31F10.13c</td>
</tr>
<tr>
<td>5</td>
<td>SPAPYUG7.04c</td>
<td>SPCC1620.14c</td>
<td>SPBC24C6.11</td>
<td>SPCC1620.14c</td>
</tr>
<tr>
<td>6</td>
<td>SPBC24C6.11</td>
<td>SPAC57A10.02</td>
<td>SPBC1734.08</td>
<td>SPBC3G6.02</td>
</tr>
<tr>
<td>7</td>
<td>SPBC1734.08</td>
<td>SPAC1556.01c</td>
<td>SPAC9.13c</td>
<td>SPAC1556.01c</td>
</tr>
<tr>
<td>8</td>
<td>SPAC23D3.09</td>
<td>SPAC30D11.13</td>
<td>SPAC23D3.09</td>
<td>SPBC821.05</td>
</tr>
<tr>
<td>9</td>
<td>SPAC17A5.11</td>
<td>SPCC970.07c</td>
<td>SPCC57A10.02</td>
<td>SPBC668.10</td>
</tr>
<tr>
<td>10</td>
<td>SPAC16A10.05c</td>
<td>SPAC23H4.12</td>
<td>SPAC16A10.05c</td>
<td>SPCC1919.15</td>
</tr>
<tr>
<td>11</td>
<td>SPC11E10.07c</td>
<td>SPCC11E10.08</td>
<td>SPCC895.07</td>
<td>SPBC23E6.09</td>
</tr>
<tr>
<td>12</td>
<td>SPBC146.13c</td>
<td>SPBC21D10.12</td>
<td>SPCP31B10.05</td>
<td>SPBC21D10.11c</td>
</tr>
<tr>
<td>13</td>
<td>SPAC821.09</td>
<td>SPCC1672.06c</td>
<td>SPCC550.12</td>
<td>SPAC22H10.07</td>
</tr>
<tr>
<td>14</td>
<td>SPBC23E6.09</td>
<td>SPAC3C7.03c</td>
<td>SPAC23H4.12</td>
<td>SPCC338.08</td>
</tr>
<tr>
<td>15</td>
<td>SPCC1682.16</td>
<td>SPBC32F12.02</td>
<td>SPBC11C11.02</td>
<td>SPBP3S2G2.08c</td>
</tr>
<tr>
<td>16</td>
<td>SPBC106.01</td>
<td>SPCC736.04c</td>
<td>SPBC1289.11</td>
<td>SPBC29A3.14c</td>
</tr>
<tr>
<td>17</td>
<td>SPCC622.16c</td>
<td>SPAC30D10.13c</td>
<td>SPCC126.02c</td>
<td>SPAC29A4.18</td>
</tr>
<tr>
<td>18</td>
<td>SPAC22H10.07</td>
<td>SPCC737.09c</td>
<td>SPBC13G1.13</td>
<td>SPBC1921.03c</td>
</tr>
<tr>
<td>19</td>
<td>SPBC13G1.13</td>
<td>SPBC56F2.01</td>
<td>SPCC188.13c</td>
<td>SPBC2F12.11c</td>
</tr>
<tr>
<td>20</td>
<td>SPBC16H5.06</td>
<td>SPBC2F12.11c</td>
<td>SPCC663.12</td>
<td>SPAB1E7.02c</td>
</tr>
<tr>
<td>21</td>
<td>SPBC29A10.05</td>
<td>SPAC16A3.07c</td>
<td>SPCC338.05c</td>
<td>SPAC17A5.16</td>
</tr>
<tr>
<td>22</td>
<td>SPBC28F2.07</td>
<td>SPAC5D6.05</td>
<td>SPAC343.11c</td>
<td>SPAC20H4.03c</td>
</tr>
<tr>
<td>23</td>
<td>SPC1827.08c</td>
<td>SPAC18G6.02c</td>
<td>SPCC1919.03c</td>
<td>SPCC1259.03</td>
</tr>
<tr>
<td>24</td>
<td>SPAC9G1.10c</td>
<td>SPAC19D5.01</td>
<td>SPAC589.02c</td>
<td>SPAC18G6.02c</td>
</tr>
<tr>
<td>25</td>
<td>SPBC33.03c</td>
<td>SPAC688.11</td>
<td>SPCC737.09c</td>
<td>SPAB1E7.06c</td>
</tr>
<tr>
<td>26</td>
<td>SPBC342.04</td>
<td>SPAC227.07c</td>
<td>SPBC106.16</td>
<td>SPAC1851.04c</td>
</tr>
<tr>
<td>27</td>
<td>SPC145.19</td>
<td>SPAC18S1.04c</td>
<td>SPAC3A11.14c</td>
<td>SPAC13G7.06</td>
</tr>
<tr>
<td>28</td>
<td>SPAC9G1.05</td>
<td>SPBC16H5.07c</td>
<td>SPAC589.02c</td>
<td>SPBC725.11c</td>
</tr>
<tr>
<td>29</td>
<td>SPCC4E9.02</td>
<td>SPAC17A5.18c</td>
<td>SPCC1450.05c</td>
<td>SPCC1739.12</td>
</tr>
<tr>
<td>30</td>
<td>SPC1794.11c</td>
<td>SPBC725.11c</td>
<td>SPAC8F11.03</td>
<td>SPCC663.01c</td>
</tr>
<tr>
<td>31</td>
<td>SPBC216.06c</td>
<td>SPAC23C11.08</td>
<td>SPAC13A11.04c</td>
<td>SPBC21C3.02c</td>
</tr>
<tr>
<td>32</td>
<td>SPC1450.05c</td>
<td>SPBC3B8.02</td>
<td>SPAC1952.07</td>
<td>SPAC1F7.09c</td>
</tr>
</tbody>
</table>
145 of the modifiers (83.8%) were orthologous to human genes, and thus are potential drug targets to modify the SMA phenotype.

To generate our candidate human disease gene networks (that we shall call network clusters) we relied on two complementary approaches. Both approaches rely on networks constructed from published protein-protein interaction data from either human (Figure 3 and Figure S3) or fission yeast (Figure 4, Figure S1 and Figure S2).

By including two types of genetic interactions, synthetic lethals or synthetic suppressors, yeast augmented network analysis (YANA), yields the most complete unbiased list of potential modifiers. In a first approach, we converted all 145 \textit{S. pombe} modifier genes to their corresponding human orthologs. All the data were combined to create network cluster diagrams. Individual datasets were delineated using a color key. (Note that we designate each interaction as a synthetic lethal or synthetic suppressor \textit{a priori}, and include them both as it is difficult to predict which of these classes will have potential therapeutic value). Using the highest confidence (0.9) experimentally confirmed protein-protein interaction data (from www.string-db.org), we used cytoscape to draw network cluster modules that include all modifier genes. We identified 22 clusters using this approach (Figure S3). We then limited the number of networks by focusing on those modifiers that shared at least two interactions with other modifiers shrinking the number of clusters to 10 (Figure 3). Within our human primary network clusters, we identified several genes that represent compelling therapeutic targets. These included GSK3 (Glycogen Synthase Kinase-3), CUL3 (Cullin-3) and NEDD4 (E3 ubiquitin ligase) along with SUMO1 (Small Ubiquitin Modifier-1) and HDAC (Histon deacetylase), a known chromatin modifier. Interestingly, GSK3 inhibitors increase SMN levels in spinal muscular atrophy patient-derived fibroblasts and mouse motor neurons\textsuperscript{28}. The HDAC inhibitor Trichostatin A increases SMN expression and survival in a mouse model of spinal muscular atrophy\textsuperscript{29}. Therefore, evidence already exists that genes within our clusters affect SMA biology.

In our second approach, we constructed networks from just the \textit{S. pombe} modifier genes using high confidence \textit{S. pombe} protein-protein interaction data from STRING. For this analysis, we identified network clusters that were expanded to include all the so-called “first-degree (1°) interacting proteins”, proteins not identified in our SGA screens but that physically interact with a genetically identified modifier (Figure S1). The rationale underlying our approach to generate larger clusters comes from the facts that the 1° neighbor may be an essential gene (and therefore it is not present in our deletion library) or may fail to be detected in our screens for other,
Figure 3. Human UBA1 primary gene interaction network. Human orthologs of yeast genes identified from four independent SGA screens were used to map the corresponding protein-protein interaction network. Only proteins that have experimentally verified data (STRING) supporting a physical interaction with at least two additional genes/proteins identified in the UBA1-SGA screens are displayed. Of particular interest is the discovery of NEDD4, CUL3, GSK3, SUMO and HDAC in distinct network clusters. Proteins of interest are highlighted by a larger node size and indicated with an arrow.
Figure 4. *S. pombe* extended UBA1 gene interaction network (extracted from Figure S1 on the basis of identifying new primary hit interactions through a 1° neighbor). Primary yeast gene hits from all four SGA screens were combined and mapped based on highest confidence experimentally confirmed protein-protein interaction data. For this extended analysis, we also included first-degree neighbors of modifiers that were not directly identified in the SGA screens. Only selected gene clusters are shown here; the complete extended network is shown in Figure S1. When extrapolated to human, the yeast extended network analysis identified Csn5 and Cul3, two modifiers also identified in our human UBA1 primary gene interaction network (boxed; see also Figure 3).
unknown reasons. This “Cluster with 1° neighbor” analysis reveals new interactions not observed in the yeast primary \textit{UBA1} gene interaction network (Figure S2). A few examples of these types of clusters are shown in Figure 4. The clusters include several genes all of which were uniquely identified in the wildtype screen, including both synthetic lethals and synthetic suppressors. Of particular interest is the cluster of proteins surrounding Skp1 (S-phase kinase associated protein-1), a protein that interacts and stabilizes F-box proteins, additional F-box proteins, and proteins involved in culin deneddylation and neddylation. In addition, we found that the RNA-binding protein Cwf2 interacts directly with three of our modifiers, and the plastin ortholog Fim1, interacts with two. Interestingly, Plastin 3 (PLS3) has been identified as a disease modifier in animal models of spinal muscular atrophy\cite{30-32}.

We then identified correlations between the extended \textit{S. pombe} network clusters (Figure 4) and the human primary network clusters (Figure 3). In both networks, we identified components of the COPS complex, a known regulator of E3 cullins, and the E3 cullin Cul3. Note that Cul3 was not present in our primary \textit{S. pombe} network (Figure S2) yet was present in the more extended network clusters that incorporate 1° neighbors; this illustrates the importance of adding 1° neighbors to our analysis of the \textit{S. pombe} data.

Together, these data suggest a prominent role for E3 ubiquitin ligases in modulating \textit{UBA1} and potentially \textit{SMN1}, making Cul3 the primary target for further analysis. The E3 ubiquitin ligase might inhibit SMN function, as ubiquitination often leads to protein degradation. If so, we reasoned that inhibition of E3 ubiquitin ligase might enhance SMN activity and thus suppress loss of function mutations in SMN1. To directly test the hypothesis that inhibition of E3 ubiquitin ligases can suppress a SMN1-spinal muscular atrophy phenotype, we turned to a vertebrate model of SMN deficiency previously reported in zebrafish\cite{33}. Knockdown of Smn in zebrafish embryos causes developmental defects in motor neuron axonal outgrowth that include truncations and abnormal branching of neurons\cite{24}. Motor neuron axon defects can be corrected by injection of mRNAs encoding wildtype human SMN\cite{25}. We first confirmed that injection of \textit{smn} morpholino (MO) into wildtype zebrafish causes severe motor axon abnormalities as compared to control uninjected embryos (Figure 5A and 5B). We then tested whether an inhibitor of E3 ubiquitin ligases (MLN4294) would

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{MLN4294 rescues the abnormal neuron outgrowth in Smn-depleted zebrafish. (A) Representative lateral views of motor axons in \textit{Tg(mnx1:GFP)} zebrafish embryos expressing GFP in motor neurons and injected with control MO and then grown in 10 or 15 \textmu M MLN4294. (B) Quantification of the effects of MLN4294 on motor axon development in zebrafish. Motor axons were scored in \textit{Tg(mnx1:GFP)} embryos injected with control MO, and subsequently (10 hrs post-injection) incubated in 10 or 15 \textmu M MLN4294. Embryos were classified as severe, moderate, mild, or no defects based on the severity of motor axon defects, and the percentage of each group is shown. Data in all graphs are represented as mean and SEM. (C) Western blot analysis of Smn and HuD protein following treatment with MLN4294. Quantification of the results is shown in Figure S5.}
\end{figure}
uninjected embryos (P=0.379). At higher concentrations of drug (20 µM), the rescue was less pronounced (Table 2) probably because higher levels of drug led to defects in development. Thus, E3 ligase inhibition can rescue neuronal defects caused by Smn protein depletion in zebrafish. In control experiments, MLN4294 failed to rescue the zebrafish mutant topped (Figure 6), a mutant defective in neuronal axon guidance. We also tested a drug that inhibits sumolyation suppress the motor axon abnormalities. Addition of the Nedd8-E1 activating enzyme inhibitor MLN4294 (that blocks cullin-RING E3 ligases), at concentrations ranging from 10 µM to 15 µM, caused a concentration-dependent reduction in the degree of abnormal motor axon branching (Figure 5A and 5B). At 15 µM inhibitor we found motor axon abnormalities that were completely suppressed: defects were not significantly different to those observed in un.injected embryos (P=0.379). At higher concentrations of drug (20 µM), the rescue was less pronounced (Table 2) probably because higher levels of drug led to defects in development. Thus, E3 ligase inhibition can rescue neuronal defects caused by Smn protein depletion in zebrafish. In control experiments, MLN4294 failed to rescue the zebrafish mutant topped (Figure 6), a mutant defective in neuronal axon guidance. We also tested a drug that inhibits sumolyation

<table>
<thead>
<tr>
<th></th>
<th>Control MO</th>
<th>smn MO</th>
<th>smn MO + 10 µM</th>
<th>smn MO + 15 µM</th>
<th>smn MO + 20 µM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe</td>
<td>0 (± 0)</td>
<td>10.1 (± 6.3)</td>
<td>3.5 (± 2.3)</td>
<td>0 (± 0)</td>
<td>2.8 (± 2.4)</td>
</tr>
<tr>
<td>Moderate</td>
<td>4.4 (± 1.7)</td>
<td>39.7 (± 1.2)</td>
<td>19.6 (± 7.2)</td>
<td>16.4 (± 5.6)</td>
<td>11.3 (± 7.1)</td>
</tr>
<tr>
<td>Mild</td>
<td>44.7 (± 6.4)</td>
<td>39 (± 7.2)</td>
<td>53.1 (± 5)</td>
<td>40.6 (± 3.4)</td>
<td>46.8 (± 9.6)</td>
</tr>
<tr>
<td>No Defects</td>
<td>50.9 (± 7.2)</td>
<td>11.2 (± 3.2)</td>
<td>23.9 (± 5.5)</td>
<td>43 (± 7.9)</td>
<td>39.1 (± 19)</td>
</tr>
<tr>
<td>n (per exp)</td>
<td>23/19/27/25</td>
<td>23/20/28/27</td>
<td>21/23/27/26</td>
<td>19/17/26/26</td>
<td>22/25/26</td>
</tr>
</tbody>
</table>

Figure 6. Treatment of zebrafish with MLN4294 fails to rescue defects in ventral motor axon guidance in the topped mutant. (A) Quantification of the effects of MLN4294 on motor axon guidance in the zebrafish topped mutant. Motor axons were scored in embryos injected with DMSO, 10 or 15 µM MLN4294. Embryos were classified as severe, moderate, mild, or no defects based on the severity of motor axon defects, and the percentage of each group is shown. Data in all graphs are represented as mean and SEM. (B) Experimental data.
corresponding to an unrelated target (SUMO) identified in our human primary network clusters (Figure 3); addition of this compound had no effect on the spinal muscular atrophy model (Figure S6) necessarily surprising since pm3, which encodes the yeast ortholog of SUMO1, has the opposite effect of cul3 when deleted. Having been identified as a synthetic suppressor rather than a synthetic lethal, perhaps compounds that activate rather than inhibit sumolyation would be beneficial as a therapeutic.

To determine how E3 ubiquitin ligases might influence Smn1 function, we examined the protein levels for both Smn and HuD. HuD is a known SMN interacting protein that binds RNAs controlling their translation and stability and functions in neural development and plasticity. We observed greater than a 2-fold increase in both Smn and Hud protein levels following treatment with MLN4294, a NEDD8 activating enzyme (NAE) inhibitor that prevents activation of E3 ligases (Figure 5C and Figure S5). It is therefore possible that rescue of the neuronal defects in our zebrafish model might involve stabilization of Smn as well as of additional proteins within the Smn network.

Growth rate data of expression of UBA1 in S. pombe

1 Data File

http://dx.doi.org/10.5256/f1000research.4188.d28505

UBA1 was expressed in S. pombe under inducing (- thiamine) and non-inducing conditions (+ thiamine) in wildtype strains (wtUBA1) and the disease-causing variant wildtype 1617UBA1 (mutUBA1), under the control of a regulatable nmt1 promoter. Growth rate data of expression of UBA1 under inducing (- thiamine) and non-inducing conditions (+ thiamine) in wtUBA1 and mutUBA1 are shown. For each condition there are six replicates.

Discussion

In this paper we introduce Y ANA, a fission yeast genetic assay that when applied to a specific human disease gene can leverage protein-protein interaction data to characterize human disease networks. From these networks, we can identify and prioritize genetic pathways likely to modify the disease-associated gene activity, and predict those genes that can be exploited as therapeutic targets. Using UBA1, a gene associated with X-linked spinal muscular atrophy (XL-SMA), we found a high degree of homology between our yeast modifiers and their corresponding human genes (>80%), underscoring the extensibility of the assay. Specifically, we identified several network clusters and high priority targets for therapeutic intervention, including GSK3 and HDAC both of which are potential therapeutic targets for SMA. In addition, we found two novel related targets Cul3 (a cullin required for E3 ubiquitin ligase activity) and NEDD4, an E3 ubiquitin ligase.

Our yeast system, with its small but complex eukaryotic genome, and complete deletion library, is unique in allowing unbiased genome-wide screening of deletions that alter human disease gene activity. Moreover, YANA can be applied to any human gene, regardless of the phenotype or availability of endogenous mutations. The number of candidate genes identified by YANA for UBA1 represents a fraction of the approximately 30,000 genes in the human genome, providing a significant enrichment of potential modifiers. Therefore, YANA offers a simple, cost-effective, and relatively rapid technology that could be applied to all human genetic diseases.

Data availability

F1000Research: Dataset 1. Growth rate data of expression of UBA1 in S. pombe, 10.5256/f1000research.4188.d28505

Author contributions

G.D. conceived the project. D.W., I.J., L.T.H performed all experiments. X.C. assisted in the bioinformatics analysis. L.B. identified the UBA1 mutation linked to XL-SMA and was co-PI on the MDA grant that supported this research. C.B. was responsible for designing the zebrafish experiments and scoring the results.

Competing interests

No competing interests were disclosed.

Grant information

This work was supported by a grant from the Muscular Dystrophy Association (MDA186435). (Christine Beattie grant support RO1NS050414 (C.E.B.) with additional support from P30NS045758). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We thank Christian Romero for his expert technical assistance. Special thanks to W. Burhans, C. Boone, B. Andrews, M. Constanzo, C. Myers, C. Nislow, L. Hartwell, J. Dallman, and P. Goldschmidt for their encouragement and discussions during the early phase of these studies.
Figure S1. *S. pombe* extended UBA1 interaction network (complete). The full protein-protein interaction network including first-degree neighbors for all fission yeast gene modifiers identified from our two independent SGA screens.
Figure S2. *S. pombe* UBA1 primary gene interaction network. Fission yeast epistatic modifiers from all four SGA datasets were combined and used to map a protein-protein interaction network using the highest confidence protein interaction data (STRING-db).

Figure S3. Complete human UBA1 primary gene interaction network representing 22 clusters. Human orthologs of yeast genes identified from four independent SGA screens were used to map the corresponding protein-protein interaction network.
Figure S4. Validation of SGA data for cul3, pub1 and gsk3. Growth curves of the 3 double mutants confirmed our original SGA data showing that the deletions of cul3 and gsk3 are synthetically lethal with OE of UBA1, and the deletion of pub1 (NEDD4) suppresses the toxicity associated with OE of wildtype UBA1. For gsk3 and cul3, the cells were washed to remove thiamine and immediately placed in the plate reader whilst for pub1, cells were first pre-induced for 22 hrs. Deletions of cul3, pub1, and gsk3 were confirmed by PCR.
Figure S5. Expression of Smn and HuD in zebrafish embryos. Quantification of western blot results shown in Figure 5C. Treatment of zebrafish embryos with MLN4294 results in approximately a 3-fold increase in both Smn and HuD levels 18 hrs following treatment.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Control MO</th>
<th>smn MO</th>
<th>smn MO + 10 μM</th>
<th>smn MO + 15 μM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLN4924 -</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>EXP #</td>
<td>1 2 3 4 1 2 3 4 1 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sev</td>
<td>0 0 0 0 0 0 0 4 7 1 4 2 1 0 0 0 0 0 0 1 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mod</td>
<td>1 1 0 2 9 8 12 10 7 6 8 3 1 4 5 3 1 4 1 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mild</td>
<td>8 12 10 11 13 9 7 8 12 10 11 9 16 8 8 8 11 12 9 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no defects</td>
<td>14 6 17 12 1 3 5 2 4 0 2 5 8 9 7 4 15 14 5 15 9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>23</td>
<td>19</td>
<td>27</td>
<td>25</td>
</tr>
</tbody>
</table>

Figure S6. Zebrafish raw data. A) MLN4924 treatment in smn MO zebrafish. B) MLN4924 treatment in topped mutant zebrafish. C) Anacardic acid treatment in smn MO zebrafish.
References


Jasper Rine  
Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA, USA

The Wiley et al manuscript describes a beautiful synthesis of contemporary genetic approaches to, with astonishing efficiency, identify lead compounds for therapeutic approaches to a serious human disease. I believe the importance of this paper stems from the applicability of the approach to the several thousand of rare human disease genes that Next-Gen sequencing will uncover in the next few years and the challenge we will have in figuring out the function of these genes and their resulting defects. This work presents a paradigm that can be broadly and usefully applied.

In detail, the authors begin with gene responsible for X-linked spinal muscular atrophy and express both the wild-type version of that human gene as well as a mutant form of that gene in S. pombe. The conceptual leap here is that progress in genetics is driven by phenotype, and this approach involving a yeast with no spine or muscles to atrophy is nevertheless an N-dimensional detector of phenotype. The study is not without a small measure of luck in that expression of the wild-type UBA1 gene caused a slow growth phenotype which the mutant did not. Hence there was something in S. pombe that could feel the impact of this protein. Given this phenotype, the authors then went to work and using the power of the synthetic genetic array approach pioneered by Boone and colleagues made a systematic set of double mutants combining the human expressed UBA1 gene with knockout alleles of a plurality of S. pombe genes. They found well over a hundred mutations that either enhanced or suppressed the growth defect of the cells expressing UBI1. Most of these have human orthologs. My hunch is that many human genes expressed in yeast will have some comparably exploitable phenotype, and time will tell.

Building on the interaction networks of S. pombe genes already established, augmenting these networks by the protein interaction networks from yeast and from human proteome studies involving these genes, and from the structure of the emerging networks, the authors deduced that an E3 ligase modulated UBA1 and made the leap that it therefore might also impact X-linked Spinal Muscular Atrophy.

Here, the awesome power of the model organism community comes into the picture as there is a zebrafish model of spinal muscular atrophy. The principle of phenologs articulated by the Marcotte group inspire the recognition of the transitive logic of how phenotypes in one organism relate to phenotypes in another. With this zebrafish model, they were able to confirm that an inhibitor of E3 ligases and of the Nedd8-E1 activating suppressed the motor axon anomalies, as predicted by the effect of mutations in S. pombe on the phenotypes of the UBA1 overexpression.

I believe this is an important paper to teach in intro graduate courses as it illustrates beautifully how
important it is to know about and embrace the many new sources of systematic genetic information and apply them broadly.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Corey Nislow
Department of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada

This paper addresses the vexing problem of how to unravel the complexity of human diseases. Such diseases often involve interactions between genes, and the question is therefore well-suited to an automated genetic approach, such as Synthetic Genetic Array technology, which was pioneered in bakers yeast and more recently adopted to another well-characterized model system, namely S. pombe. In this study the authors examine conserved genetic interactions between a fission yeast genes and a human query gene (both wildtype and loss-of-function mutants) then combine their experimental data with bioinformatics to characterize the network of genetic interactions and protein-protein interactions. They then leverage this network to identify interacting partners that comprise novel therapeutic targets.

The example presented in this paper is the human UBA1 gene which encodes a gene that, when mutated, results in human X-linked spinal muscular atrophy disease. They perform whole genome genetic interaction mapping for digenetic interactions with both wild type and mutant versions of the gene, build a network with the observed interactions, and then expand that primary network by adding the known partners that interact with members of the primary network. They then turned to the zebrafish vertebrate model system to validate their newly identified therapeutic targets.

This study adds another tool to the armamentarium of systems biology approaches with which the complexity of human diseases can be dissected. Because the system is new in the sense that fission yeast has been largely unexplored (compared to its budding yeast relative from which it has diverged by several hundred million years) the observations on fission yeast, when combined with those from other model systems, provide an attractive complementary approach to understanding complex human biology.

Given that this study builds on earlier approaches, the authors should cite the previously published studies on genetic interactions, chemical genetic interactions and protein-protein interactions using budding yeast, fly and other models. Similarly, the comprehensive gene expression profiling approaches used in cultured mammalian cells as well as more recent whole genome RNA interference and CRISPR-based disruption screens should be cited.

From a methodological perspective, in the interaction maps presented in Figures 3 and 4 it would be useful to describe or cite the methods used to define the links between genes. In the Results, where the authors describe the results of the different screens which include wild type and mutant human protein interactions as well as primary and secondary interactions it would be instructive to provide a graphic or flow chart that depicts how all the key pieces of data from each independent screen in fission yeast (and their analysis) were collapsed and presented together. It would also be instructive to provide background
information or appropriate citations on the two compounds used in the zebrafish experiments, specifically their known or suspected mechanisms of action as well as the data to support these mechanisms.

In the Discussion, the authors have room to put the YANA screening platform into context, comparing it to other orthologous genome-wide approaches.

A key innovation in this study arises from the unique approach used in analyzing the data. Specifically, the conserved yeast genes that modify the phenotype induced by expressing a human gene of interest are assembled into human disease gene networks which are then augmented by additional protein-protein interaction data. Augmenting traditional network analysis in this manner may permit the identification of new therapeutic targets and target pathways. In summary, this comprehensive, genome-wide work represents an important addition to the systems biology toolkit to understand the complexities of human biology and the impact of this and subsequent data sets will surely increase as additional screens are added to the compendium.

Minor points

- Western blot and growth curve: "or eight generations" is a typo, also the type of media the cells were grown in "exponentially for 16 hours" should be listed.

- In Figure 2 it is not clear which band the tubulin arrow is pointing to- the loading control bands should be indicated directly on the blot/gel. Also, it is apparent from this loading control that the amount of UPA1 expressed is extremely high, the authors should comment on the fact that the overexpressing lanes are actually UNDERloaded relative to the non-expressing lanes.

- Third paragraph of the Results: the software programme cytoscape should include a reference.

- Third paragraph of the Results: histone is misspelled "histon".

- Sixth paragraph of the Results: the authors should clearly distinguish between the published data in zebrafish and their new data collected in this study

- Sixth paragraph of the Results: as mentioned be inhibitor's mechanism of action should be described in more detail

- Second paragraph of the discussion: approximately 30,000 genes in the human genome should be cited and probably corrected to the conventional number of 21,000.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

**Competing Interests:** I was involved in some of the preliminary discussions with the authors about this work in the early phases when these studies were being conceived. However, I can confirm that I was not involved in carrying out any of the research nor in the writing of the final article.