Genetic Basis of Variation in Ubiquitin-Proteasome System Activity

Mahlon A. Collins[1], Gemechu Mekonnen[1,2], and Frank W. Albert[1]

[1]Department of Genetics, Cell Biology, and Development
University of Minnesota
Minneapolis, MN, U.S.A.

[2]Current Address:
Department of Biology
Johns Hopkins University
Baltimore, MD, U.S.A.

Co-Corresponding Authors:
Mahlon A. Collins (mahlon@umn.edu)
Frank W. Albert (falbert@umn.edu)

Abstract

The ubiquitin-proteasome system (UPS) is the cell's primary pathway for targeted protein degradation. Although the molecular mechanisms controlling UPS activity are well-characterized, we have almost no knowledge of how these mechanisms are shaped by heritable genetic variation. To address this limitation, we developed an approach that combines fluorescent UPS activity reporters with a statistically powerful genetic mapping framework to comprehensively characterize genetic influences on UPS activity in the yeast *Saccharomyces cerevisiae*. We applied this approach to substrates of the UPS N-end rule, which relates a protein's degradation rate to the identity of its N-terminal amino acid ("N-degron") through the Arg/N-end and Ac/N-end pathways. Genetic influences on UPS activity were numerous and complex, comprising 149 loci influencing UPS activity across the 20 N-degrons. Many loci specifically affected individual pathways or degrons and multiple loci exerted divergent effects on distinct UPS pathways. One Arg/N-end pathway-specific locus resulted from multiple causal variants in the promoter, open reading frame, and terminator of the *UBR1* E3 ubiquitin ligase gene. These variants differentially affected substrates bound by the Type 1 and Type 2 recognition sites of Ubr1p. Collectively, our results provide the first systematic characterization of genetic influences on UPS activity and a generalizable approach for mapping genetic effects on protein degradation with high statistical power and quantitative precision.

Introduction

Protein degradation by the ubiquitin-proteasome system (UPS) is an essential biological process that regulates the abundance of cellular proteins and removes damaged and misfolded proteins from cells[1–3]. Through these actions, UPS protein degradation influences multiple aspects of cellular physiology, including energy metabolism[4–7], cell signaling cascades[8–10], and stress responses[11–13]. Aberrant UPS activity adversely affects multiple cellular functions and contributes to a diverse array of human diseases, including cancers[14,15] and neurodegenerative[16–18], immune[19,20], and metabolic disorders[21,22]. To what extent both physiological and pathological variation in UPS activity is driven by natural genetic variation is almost entirely unknown, aside from a handful of limited examples. Genome-wide association studies have linked variation in UPS genes to the risk for several diseases, but the vast majority of such studies have not characterized how these risk loci influence UPS activity[23–25]. Missense mutations in UPS genes cause fatal, incurable syndromes[26–28]. However, such mutations often entirely ablate the function of the associated gene, suggesting they represent only one extreme of a continuum of potential variant effects on UPS activity. Beyond these examples, we do not know which DNA sequence variants in a population affect UPS activity, their magnitude of effect, or the molecular mechanism(s) by which they influence UPS activity. Consequently, we do not know how such variants contribute to the genetic basis of complex organismal and cellular traits regulated by the UPS, including the many diseases marked by aberrant UPS activity[20].

Our limited understanding of how genetic variation influences UPS activity stands in stark contrast to the well-characterized molecular mechanisms of UPS function. UPS protein degradation comprises a series of enzymatic reactions that carry out the steps of recognizing and marking proteins for degradation, trafficking marked proteins to the proteasome, and substrate unfolding and degradation by the proteasome[1,3]. This process begins when E3 ubiquitin ligases recognize degradation-promoting signal sequences, termed degrons[29], in cellular proteins. When an E3 ligase recognizes its cognate degron, the small protein ubiquitin is covalently attached to the substrate protein, marking it for proteasomal degradation[30,31]. The proteasome then binds, unfolds, and degrades marked substrates to individual amino acids or small peptides[32,33]. The activity of each of these reactions is tuned by a diverse array of regulatory mechanisms, many of which are specific to individual UPS pathways. These include transcriptional programs that control the abundance of UPS gene products[34,35], post-translational modifications of UPS components that modify their activity[36,37], and signaling cascades that alter UPS activity based on the metabolic state of the cell[38,39]. The large number and diversity of regulatory mechanisms controlling UPS protein degradation creates many potential targets through which genetic variation may shape UPS activity.

Comprehensively characterizing how genetic variation shapes UPS activity thus requires assays that can be readily adapted to multiple UPS pathways, degrons, and substrates. Genetic mapping with controlled crosses of model organisms has been productively used to understand the genetic basis of variation in a variety of cellular and organismal traits[40–42]. However, technical challenges have thus far precluded using such approaches to characterize genetic influences on UPS activity. In particular, to attain high statistical power, genetic mapping methods require assays that can measure a trait of interest in large numbers of individuals, often in the thousands[41,43–45]. At this scale, many assays of protein degradation would be prohibitively labor- and time-intensive[46]. Several high-throughput techniques exist for measuring UPS activity[46–48]. However, to what extent these methods can be deployed in genetically diverse mapping populations without confounding from differences in reporter expression is unknown. Recent advances in the design of fluorescent timers have the potential to overcome this limitation by providing readouts of UPS activity that are unaffected by differences in reporter expression[49]. In particular, the tandem fluorescent timer (TFT) design has recently been adapted to provide high-throughput measurements of a variety of UPS pathways and substrates[49–53]. This system could, in principle, enable genetic mapping of UPS activity.

Here, we leveraged recent advances in the design of TFTs and genetic mapping to systematically characterize the genetic basis of variation in UPS activity. We built a series of TFTs that provide high-throughput, quantitative measurements of the Arg/N-end and Ac/N-end UPS pathways comprising the UPS N-end rule. We used these reporters to perform bulk segregant genetic mapping in a cross of two strains of the yeast *Saccharomyces cerevisiae*. Our results revealed considerable complexity in the genetic influences on UPS activity. Individual UPS degrons are affected by multiple loci, many loci act in a pathway-specific manner, and individual loci influencing UPS activity can contain multiple causal variants in the same gene that act through distinct molecular mechanisms. More broadly, our work introduces a generalizable strategy for mapping genetic influences on the UPS and protein degradation with high statistical power and quantitative precision.

Results

Ubiquitin-Proteasome System Activity Reporters

To map genetic influences on the UPS, we first built and characterized a series of TFT reporters capable of measuring UPS activity with high-throughput and quantitative precision in living cells. TFTs are linear fusions of two fluorescent proteins with distinct spectral profiles and maturation kinetics[48,54]. The most commonly implemented TFT consists of a faster-maturing green fluorescent protein (GFP) and a slower-maturing red fluorescent protein (RFP)[48,49,54] (Figure 1A). If

the degradation rate of a TFT-tagged substrate is faster than the maturation rate of the RFP, then the RFP / GFP ratio (hereafter, "TFT ratio") is inversely proportional to the substrate's degradation rate[48,49,54] (Figure 1B). The TFT ratio is also independent of reporter expression[49], allowing high-throughput measurements of UPS activity in genetically diverse populations of cells.

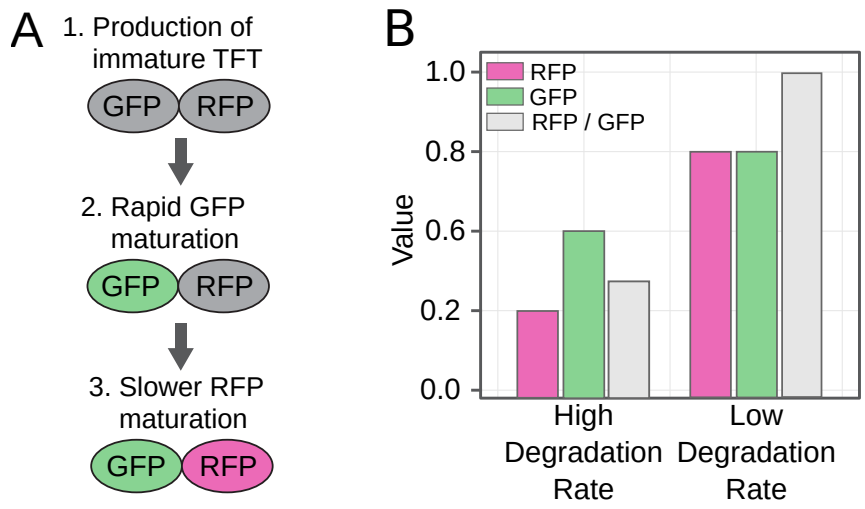


Figure 1: Tandem fluorescent timer (TFT) overview. A. Schematic of the production and maturation of a TFT. B. Simulated data showing the use of a TFT's RFP/GFP ratio ("TFT ratio") to measure UPS protein degradation activity.

To understand how genetic variation affects UPS activity, we focused on the well-characterized UPS N-end rule pathway. The N-end rule relates a protein's degradation rate to the identity of its N-terminal amino acid (hereafter, "N-degron")[33,55,56]. In the yeast *Saccharomyces cerevisiae*, the N-end rule pathway can be divided into two primary branches, the Arg/N-end and Ac/N-end pathways[33,55]. Arg/N-end substrates are recognized and bound by the E3 ligase Ubr1p[57]. Ubr1p has two binding sites, which recognize distinct classes of Arg/N-end degrons. The Type 1 binding site recognizes basic N-degrons, while the Type 2 binding site recognizes bulky hydrophobic N-degrons[33] (Figure 2A and D). The other primary branch of the N-end rule pathway is the Ac/N-end pathway, whose substrates are recognized and bound by the E3 ligase Doa10p[58]. Doa10p binds substrates with acetylated N-terminal amino acids, as well as unacetylated G and P residues (Figure 2A and D)[33,58]. We reasoned that the diversity of degradation signals encompassed in the N-end rule pathway would maximize our ability to identify genetic variation affecting UPS activity and that the well-characterized effectors of the N-end rule pathway would aid in defining the molecular mechanisms of variant effects on UPS activity.

Analysis of N-end Rule TFTs

We built TFTs bearing each of the 20 possible N-degrons of the N-end rule pathway. To do so, we attached TFTs to the C-terminus of a previously-characterized peptide sequence that engages the N-end rule pathway[33,56,58]. This peptide is unstructured, lacks internal degrons[59], contains internal lysine residues that are efficiently ubiquitinated[56,58], and has previously been used to characterize the function of the N-end rule pathway[56–58,60]. To generate each of the 20 possible N-degrons, we used the ubiquitin fusion technique, in which a ubiquitin moiety is placed immediately upstream of the desired N-terminal amino acid[60]. After the construct is translated, ubiquitin hydrolases cleave the ubiquitin molecule, exposing the desired N-terminal amino acid, which functions as an N-degron[60] (Figure 2A). We devised a generalized approach for integrating N-end rule TFTs into the yeast genome at a defined genomic location (Supplementary Figure 1) and built strains harboring all 20 possible N-degron TFTs.

We used multiple strains and two genetic backgrounds to characterize our N-end rule TFTs. Our genetic mapping strains are a laboratory strain closely related to the S288C strain ("BY") and a wild vineyard strain ("RM"). These strains have previously been used to map the genetic basis of a variety of cellular and organismal traits[43,61,62]. We also constructed a series of reporter control strains by deleting individual UPS genes from the BY strain. Specifically, we built strains lacking the Arg/N-end E3 ligase *UBR1*[57] ("BY *ubr1*Δ"), the Ac/N-end E3 ligase *DOA10*[58] ("BY *doa10*Δ"), or the proteasome gene transcription factor *RPN4*[34,35] ("BY *rpn4*Δ"). Deleting *UBR1* is expected to stabilize Arg/N-end reporters and deleting *DOA10* is expected to stabilize Ac/N-end reporters[50,57,58] (Figure 2A). Because deleting *RPN4* leads to reduced numbers of proteasomes and reduced UPS activity[35], all reporters are expected to be stabilized in BY *rpn4*Δ. We built BY, RM, BY *rpn4*Δ, BY *ubr1*Δ, and BY *doa10*Δ strains harboring each of our 20 N-end rule TFTs.

We characterized UPS activity towards each N-end rule TFT in our strains by flow cytometry. Theoretical and empirical observations indicate that when a TFT's degradation rate is slower than the maturation rate of its RFP, the TFT's $\log_2$ RFP/GFP ratio will be approximately 0 when the fluorescence output of the two fluorophores is equivalent[49,50,54]. Consistent with out expectations, we observed that the $\log_2$ RFP/GFP ratio for Arg/N-end TFTs in BY *ubr1*Δ was approximately 0 and significantly greater than the corresponding $\log_2$ RFP/GFP ratio in our wild-type BY and RM strains (Figure 2B/C/D, Supplementary Figure 2, Supplementary Table 1). Deleting *DOA10* likewise produced a significant stabilization of our set of Ac/N-end TFTs relative to BY and RM (Figure 2B/C/D, Supplementary Figure 2, Supplementary Table 1). Deleting *RPN4* significantly stabilized TFTs from both the Arg/N-end and Ac/N-end pathways relative to BY and RM (Figure 2B/C/D, Supplementary Figure 2, Supplementary Table 1). We observed that the proline N-end

TFT was only partially stabilized in BY *doa10*Δ (Supplementary Figure 3), consistent with previous results[60,63]. Specifically, when followed by a proline residue, ubiquitin is inefficiently cleaved in the ubiquitin-fusion technique[33,60]. Consequently, the proline N-end TFT simultaneously measures the degradation rate of the proline N-degron and the activity of the ubiquitin-fusion degradation pathway[64]. Taken together, these results show that our reporters provide sensitive, pathway-specific, and quantitative measurements of UPS activity.

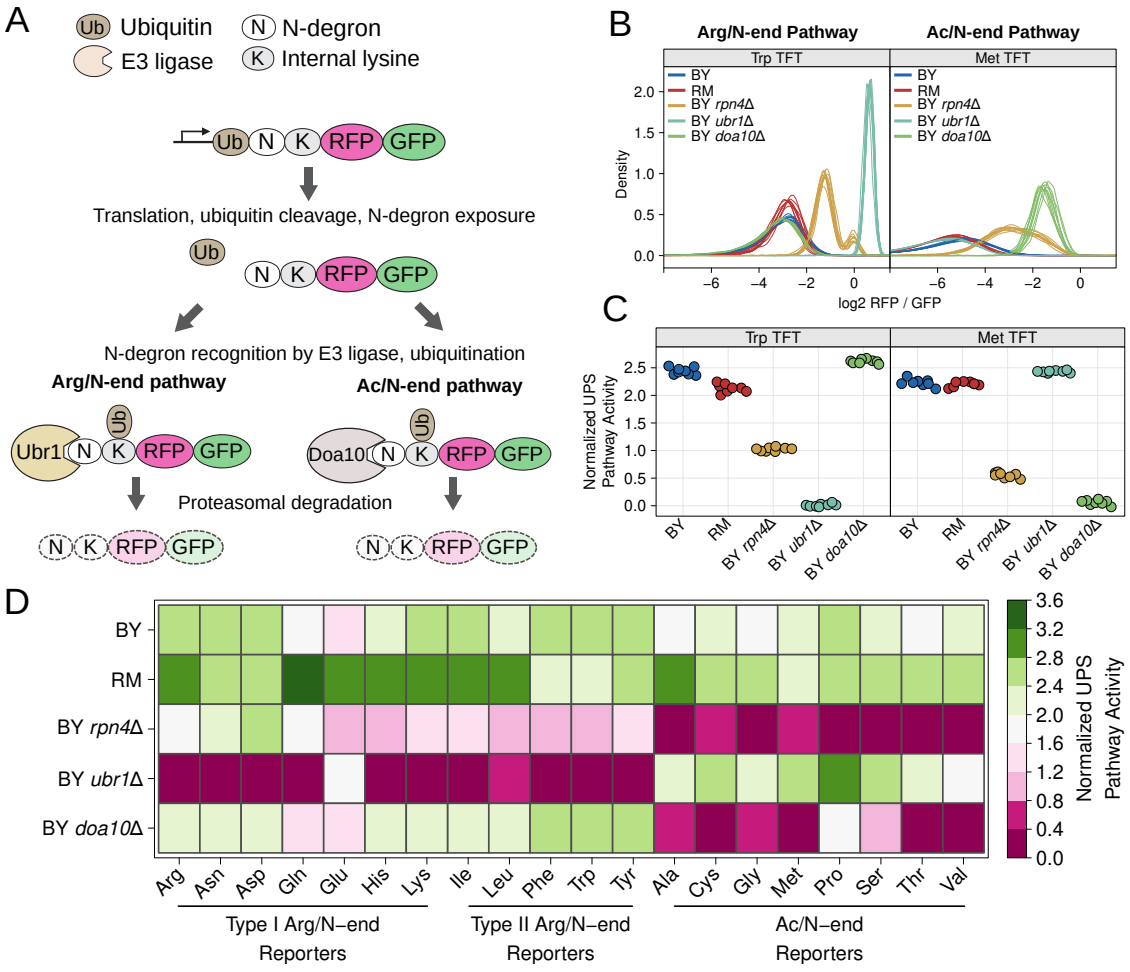


Figure 2: Design and analysis of UPS N-end rule TFTs. A. Schematic of the production and degradation of TFTs according to the UPS N-end rule. B. Density plots of the $\log_2$ TFT ratio from 10,000 cells for each of 8 independent biological replicates per strain per reporter for example Arg/N-end and Ac/N-end reporters. "BY" and "RM" are the strains used for genetic mapping. "BY *rpn4*Δ", "BY *ubr1*Δ ", and "BY *doa10*Δ" are control strains derived from BY used to characterize each TFT. C. The median of each biological replicate in B. was extracted, scaled, normalized, and plotted as a stripchart for the Trp and Met example N-degron TFTs. D. Heatmap summarizing the results for all strains and N-degrons using data generated as in C.

We also compared the degradation rate of each N-degron in the BY and RM strains to determine their overall UPS activity levels. RM had significantly higher UPS activity than BY for 9 of 12 Arg/N-end TFTs and 6 of 8 Ac/N-end TFTs (Figure 2D, Supplementary Figure 3, Supplementary Table 1). The set of N-degrons for which RM did not have significantly higher UPS activity included the Ac/N-degrons Met and Pro and the Arg/N-degrons Phe, Trp, and Tyr (Figure 2D, Supplementary Figure 3, Supplementary Table 1). The degradation of the Phe, Trp, and Tyr N-degrons was significantly higher in BY than RM, while the degradation rate of Met and Pro N-degrons were equivalent between the two strains (Figure 2D, Supplementary Figure 3, Supplementary Table 1). Overall, RM has higher UPS activity than BY, suggesting that individual genetic differences can create heritable variation in UPS activity. The observation that BY had higher UPS activity for a subset of N-degrons also raises the possibility that genetic effects on UPS could be specific to an individual N-degron.

UPS Quantitative Trait Locus Mapping by Bulk Segregant Analysis

We mapped quantitative trait loci (QTLs) for UPS activity using bulk segregant analysis, a method that achieves statistical power by comparing individuals with extreme phenotypes selected from a large population[44,65]. We created populations of meiotically recombined haploid yeast cells ("segregants") derived from mating TFT-containing BY strains with RM using a modified synthetic genetic array methodology[66,67] (Figure 3A). We collected pools of 20,000 cells from the 2% tails of the UPS activity distribution in our segregant populations using fluorescence-activated cell sorting (FACS) (Figure 3, B/C). Whole-genome sequencing was then used to determine the allele frequency difference between the high and low UPS activity pools at each DNA variant. At loci linked to UPS activity (QTLs), the allele frequencies will be significantly different between pools, while at unlinked loci the allele frequency difference will be, on average, 0. We called significant QTLs using an empirically-derived null distribution of the logarithm of the odds (LOD; see "Methods") and set the QTL significance threshold to a LOD score of 4.5, which resulted in a 0.5% false discovery rate (FDR). We further filtered our list of QTLs by retaining only QTLs detected in both of two independent biological replicates. Replicating QTLs were defined as QTLs whose peak positions were within 100 kb of each other on the same chromosome and that had the same direction of allele frequency difference in both biological replicates (Figure 4A). The complete set of replicating QTLs is found in Supplementary Table 2. The full list of QTLs, including those that did not replicate in both biological replicates is found in Supplementary Table 3. The full set of plotted allele frequency differences and LOD traces is found in Supplementary File 1.

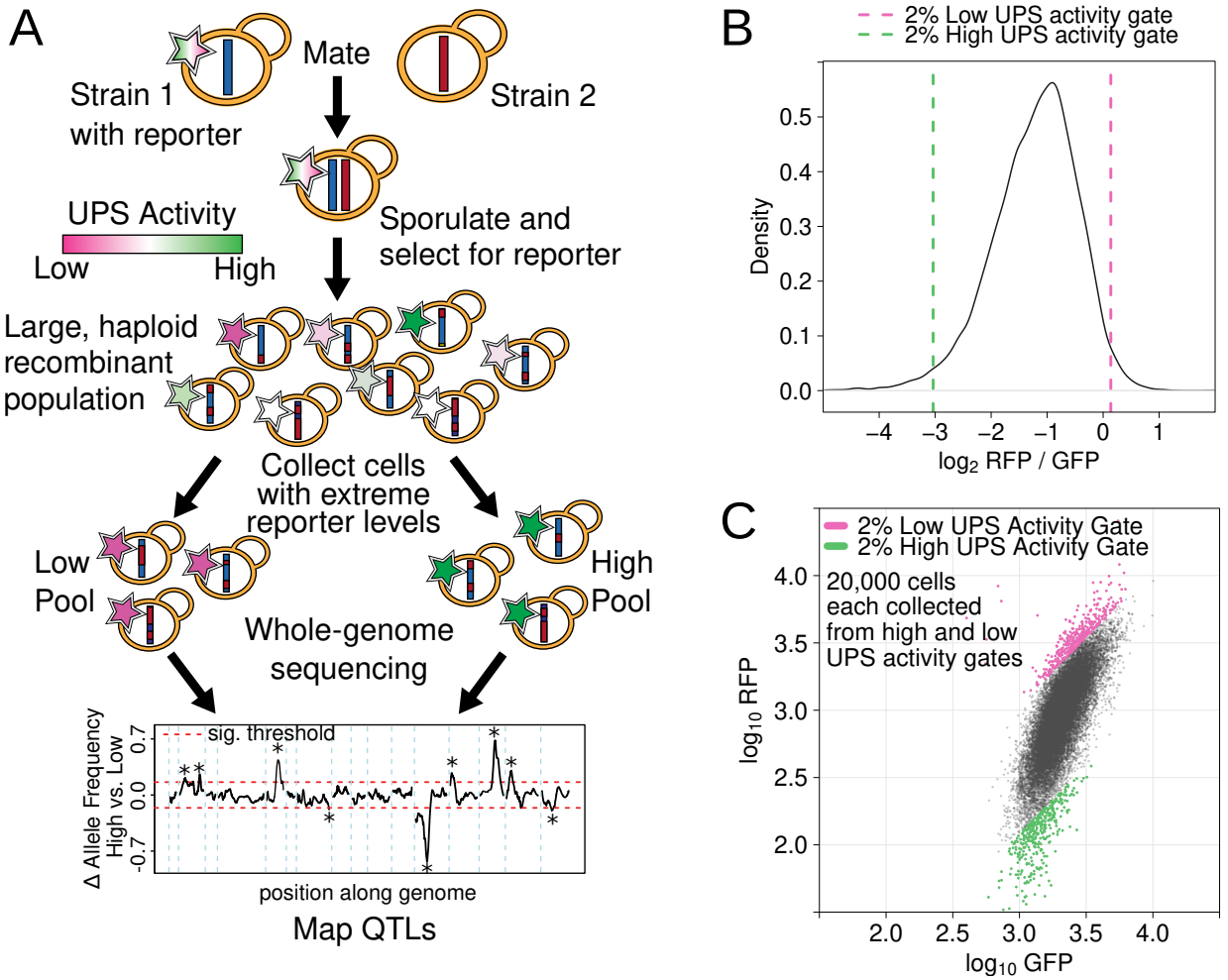


Figure 3: Bulk segregant analysis QTL mapping overview. A. Schematic of the approach for mapping UPS activity QTLs. B. Density plot of the UPS activity distribution for a segregant population. Dashed vertical lines show the gates that were used to collect cells. C. Backplot of the cells in the gates in B. onto a scatter plot of GFP (x axis) and RFP (y axis).

Global Analysis of UPS Activity QTLs

We identified 149 UPS activity QTLs across the set of 20 N-degron TFTs (Figure 4B, Supplementary Table 2). The number of QTLs per reporter ranged from 1 (for the Ile TFT) to 15 (for the Ala TFT) with a median of 7. As expected, replicating QTLs had significantly higher LOD scores (t-test p = 4e-14) and significantly greater absolute magnitudes of allele frequency differences (t-test p = 8e-30) than non-replicating QTLs (Supplementary Figure 4). Our results show that UPS activity is a genetically complex trait, shaped by many loci throughout the genome. The genetic architecture of UPS activity is characterized by a continuous distribution containing many loci of small effect and few loci of large effect (Figure 4A/B, Supplementary Figure 4).

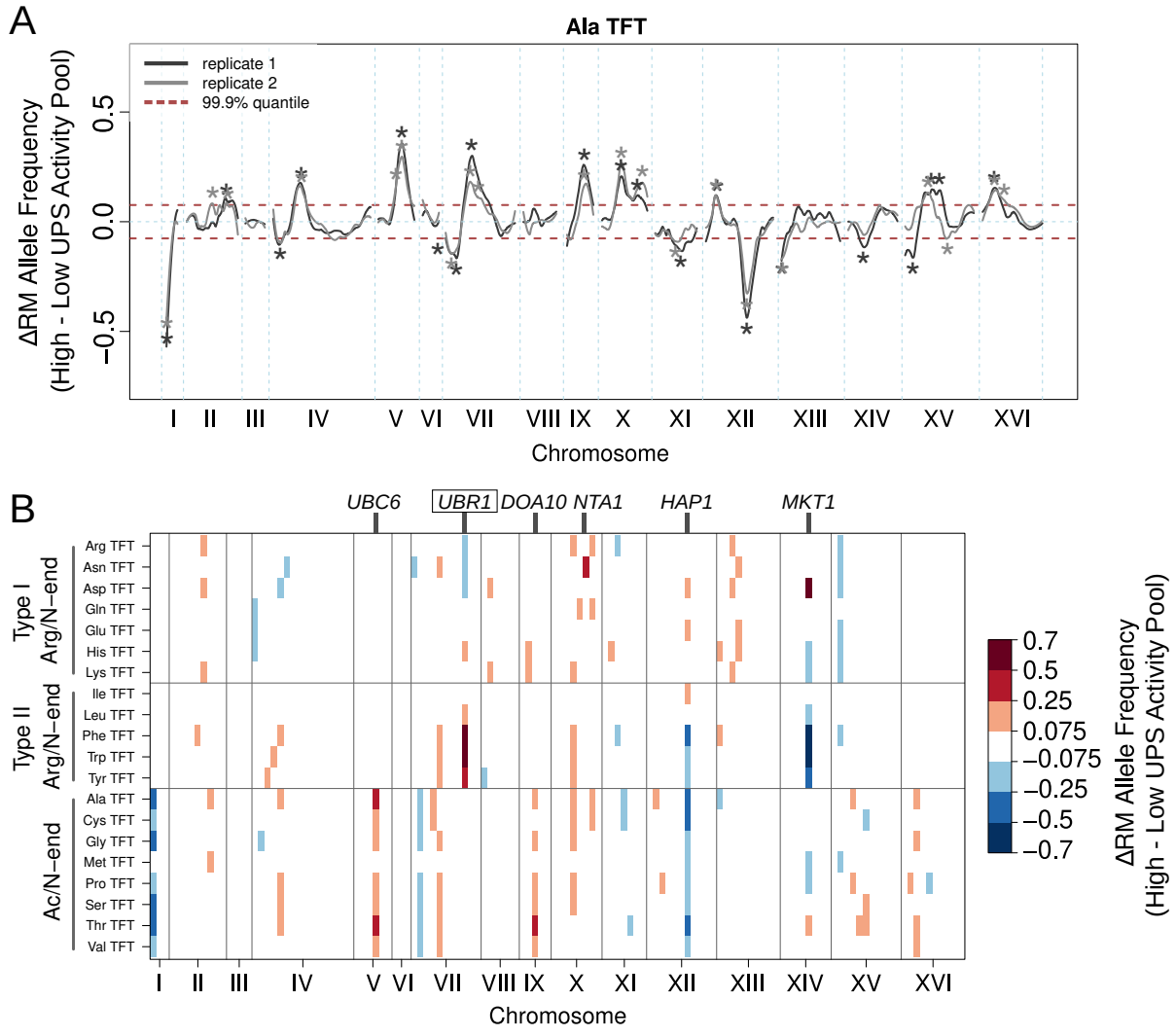


Figure 4: UPS QTL Mapping Results. A. Results from the alanine N-degron TFT are shown as an example of the results and reproducibility of the method. Asterisks denote QTLs. B. The heatmap shows the QTL mapping results for all 20 N-degron reporters. Colored blocks denote QTLs detected in each of two biological replicates, which are colored according to the direction and magnitude of the allele frequency difference. Validated (enclosed in a box) and candidate (unboxed) causal genes for select QTLs are annotated above the plot.

Analysis of the set of UPS QTLs revealed several patterns. First, we observed that the RM allele was associated with higher UPS activity in the majority of our UPS QTLs. Across all reporters, the RM allele was associated with higher UPS activity in 89 out of 149 UPS activity QTLs (60%, Supplementary Figure 5). This fraction was significantly different than the 0.5 value expected by chance (binomial test $p = 0.021$). We plotted the histogram of allele frequency differences and observed an especially strong enrichment for RM alleles associated with higher UPS activity in the range of 0.2 to 0.3, where there were approximately twice as many QTLs than the corresponding

range of -0.2 to -0.3 (Supplementary Figure 5). These results are consistent with our flow cytometry results, where we observed that RM had higher UPS activity for 15 of 20 N-degrons (Figure 2D, Supplementary Table 1), suggesting that the QTLs we have mapped underlie a substantial portion of the heritable UPS activity difference between BY and RM.

The number and patterns of QTLs differed between the Ac/N-end and Arg/N-end pathways. The Ac/N-end pathway had a significantly higher median number of QTLs per reporter than the Arg/N-end pathway (9 versus 7, respectively, Wilcoxon test $p = 0.021$). A notable difference in the patterns of QTLs between pathways was the presence of several large effect QTLs for the Arg/N-end pathway but not for the Ac/N-end pathway (Figure 4B, Supplementary Figure 5). The QTLs of large effect in the Arg/N-end pathway were found in multiple reporters on chromosomes VII and XIV.

We then evaluated the extent to which individual QTLs were shared across multiple reporters. To do so, we divided each chromosome into adjacent 100 kb bins. We considered a QTL to be shared between reporters if the peak position for two or more QTLs were within the same bin. We observed that many QTLs were unique to an individual N-degron (Figure 4B, Supplementary Figure 6), highlighting the complexity of genetic influences on N-end rule pathways. There was relatively little sharing of QTLs within the Arg/N-end pathway, for which only one QTL affected the majority of reporters (Figure 4B, Supplementary Figure 6). By contrast, the Ac/N-end reporters tended to share more of their QTLs (Figure 4B, Supplementary Figure 6). In particular, a QTL on chromosome XII affected all Ac/N-end reporters and Ac/N-end-specific QTLs on chromosomes I, V, and VII affected 7 of 8 Ac/N-end reporters (Figure 4B). These results suggest that genetic influences on the degradation of Ac/N-end rule substrates may act more broadly than those of the Arg/N-end pathway. This notion is consistent with the molecular mechanisms that generate Ac/N-degrons, which share several molecular events that create functional Ac/N-end degrons[33,58]. By contrast, the mechanisms that generate Arg/N-degrons are less general[33], consistent with the largely N-degron-specific QTL architectures we observe for this pathway. Taken together, our QTL results suggest that individual UPS pathways are shaped by distinct, complex genetic architectures and that genetic influences on UPS activity are often specific to an individual pathway.

Analysis of Individual UPS Activity QTLs

We next examined individual QTLs to better understand potential molecular mechanisms of variant effects on UPS activity. We first sought to identify QTLs resulting from genetic variation in UPS genes, reasoning that pathway-specific QTLs could result from variants in genes encoding components of the Arg/N-end or Ac/N-end pathways. A QTL on chromosome VII was detected for 8

of 12 Arg/N-end TFTs. The peak position for the QTL occurred at 861,950 bp, placing it in the *UBR1* E3 ubiquitin ligase gene. Arg/N-degrons can be divided into basic and bulky hydrophobic degrons, which are recognized and bound by the Ubr1p Type 1 and Type 2 sites, respectively. The RM allele of this QTL is associated with lower UPS activity for Type 1 N-degrons, with the exception of the His N-degron, and higher activity for Type 2 N-degrons (Figure 4B). The QTL's effect size, as measure by the allele frequency difference values, were among the highest in our set of QTLs.

A QTL on chromosome V was specific to the Ac/N-end pathway and was detected for 7 of 8 Ac/N-end reporters (Figure 4B). For all 7 reporters, the RM allele was associated with higher UPS activity. The QTL interval spans a region containing *UBC6*, the E2 ligase for *DOA10*[68,69]. The RM allele of *UBC6* contains a 3 bp deletion in the gene's promoter and a missense variant that exchanges an aspartic acid residue with a charged, bulky side chain for a glycine residue with a small, uncharged side chain at amino acid 229 in the protein. An expression QTL ("eQTL") in this same region influences the abundance of the *UBC6* mRNA[45], suggesting that this QTL may shape Ac/N-end pathway activity through effects on *UBC6* expression. An Ac/N-end QTL on chromosome IX detected for 6 of 8 Ac/N-end TFTs was centered on the *DOA10* gene. The RM allele of this QTL was associated with higher UPS activity in all 6 TFTs for which the QTL was detected. No local QTLs for *DOA10*[45] have been found, but the RM *DOA10* allele contains 3 missense variants. The other Ac/N-end specific QTLs were located on chromosomes V and VII and their intervals did not contain Ac/N-end-specific recognition components or binding factors.

A QTL on chromosome X was found only for the Asn N-degron of the Arg/N-end pathway. The QTL peak is centered on the *NTA1* gene, which encodes an amidase that converts N-terminal Asn and Gln residues to Asp and Glu residues, respectively. This conversion is necessary for the recognition and degradation of substrates with N-terminal Asn and Gln residues by the Arg/N-end pathway. The RM allele of the *NTA1* locus is associated with higher degradation of Asn N-degrons. RM *NTA1* contains two missense variants near the proton donor active site, D111E and E129G. *NTA1* is not affected by a local eQTL, making these missense variants strong candidate causal nucleotides. More broadly, these results suggest that genetic variation can influence the full sequence of individual molecular events involved in the processing, recognition, ubiquitination, and degradation of a given UPS substrate.

The BY/RM cross has been extensively used for genetic mapping of a variety of complex cellular and organismal traits, including gene expression at the mRNA and protein levels[42–45,61,62,70–73]. We examined the overlap of our UPS activity QTLs with QTLs previously identified in BY/RM genetic mapping studies. The most frequently observed UPS QTL was detected for 8 of 8 Ac/N-end and 6 of 12 Arg/N-end TFTs and was located on chromosome XII (Figure 4B). The average peak

position of the QTL was 657,046 bp, placing it in the immediate vicinity of a Ty1 element found in the *HAP1* transcription factor gene in the BY strain[74]. The Ty1 insertion in *HAP1* exerts strongly pleiotropic effects on gene expression, altering the expression of 3,755 genes[45]. Similarly, a QTL on chromosome XIV affected 10 of our 20 N-degron TFTs and was not specific to an individual UPS pathway. The average peak position of the chromosome XIV QTL was 465,450 bp, which is located in the *MKT1* gene. *MKT1* encodes a poorly-understood, multi-functional RNA binding protein involved in the post-transcriptional regulation of gene expression[12,75,76]. Variation in *MKT1* is highly pleiotropic, and the *MKT1* locus has been shown to influence the expression of 4,550 genes[45]. Thus, while *HAP1* and *MKT1* are promising candidate causal genes for their associated QTLs, the molecular mechanisms by which they influence UPS activity are likely complex and indirect.

Fine-Mapping of the *UBR1* QTL

Our analyses showed a prominent, pathway-specific role of a QTL on chromosome VII in shaping UPS activity. We selected this QTL for further molecular dissection. To test if the QTL is caused by variation in the *UBR1* gene located at its peak, we used CRISPR-Swap[77] to engineer BY strains with alternative *UBR1* alleles (Table 4, Supplementary Table 4). The full RM allele (comprising the promoter, open-reading frame [ORF], and terminator) significantly altered the degradation of all tested N-degrons (Figure 5A/B, Supplementary Table 6). Thus, *UBR1* is the causal gene at this locus.

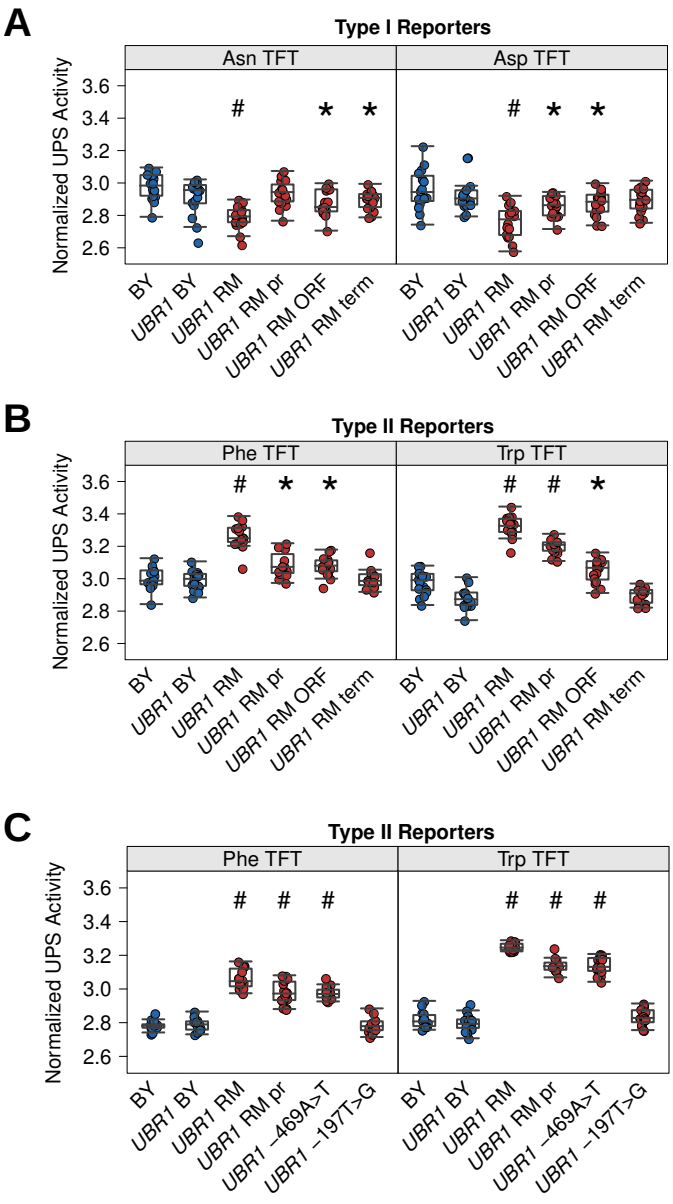


Figure 5: Fine-Mapping of the *UBR1* QTL. BY strains were engineered to carry alternative *UBR1* alleles and their UPS activity towards the indicated Arg/N-degrons was characterized by flow cytometry. Individual circles show the median of 10,000 cells for each of 16 independent biological replicates per strain per N-degron. "BY" is the wild-type strain. "*UBR1* BY" is a negative control strain which underwent allelic editing, but carries the BY *UBR1* allele. The remaining strains contain RM *UBR1* alleles as indicated. "pr" = "promoter", "term" = terminator. A. Results for the Type I Asn and Asp Arg/N-degrons. B. Results for the Type II Phe and Trp Arg/N-degrons. C. Results for the Type II Phe and Trp Arg/N-degrons with additional strains harboring single nucleotide variants in the *UBR1* promoter as indicated. In all plots, "*" = $0.05 > p >$ 1e-6 and "#"= $p <$ 1e-6 for Tukey HSD test of the indicated strain versus BY.

UBR1 contains 2 promoter, 8 missense, and 8 terminator variants between BY and RM (Supplementary Table 4). We engineered BY strains carrying chimeric *UBR1* alleles to assess the effects of variants in the *UBR1* promoter, ORF, and terminator on Arg/N-end pathway activity. Because the chromosome VII QTL exerted divergent effects on Type 1 and 2 Arg/N-end substrates, we tested the effects of RM *UBR1* alleles on the Type 1 N-degrons Asn and Asp and the Type 2 N-degrons Phe and Trp.

Variants in the *UBR1* promoter significantly decreased the degradation rate of the Asp N-degron (Figure 5A), did not affect the degradation rate of the Asn N-degron (Figure 5A), and significantly increased the degradation rate of the Phe and Trp N-degrons (Figure 5B, Supplementary Table 5). In contrast, the RM *UBR1* ORF significantly decreased the degradation rate of both the Asn and Asp N-degrons (Figure 5A, Supplementary Table 5). While both the promoter and the ORF variants significantly increased the degradation rate of the Phe and Trp N-degrons (Figure 5B), the effect of the promoter variants on the Trp N-degron were markedly stronger than those caused by the ORF variants (Figure 5B). The RM *UBR1* terminator significantly decreased the degradation of the Asn N-degron without affecting any of the other N-degrons tested (Figure 5A/B). Thus, the *UBR1* QTL contains multiple causal variants in different regions of the gene that exert unique effects on the activity of the Arg/N-end pathway through distinct molecular mechanisms.

To identify the specific causal variant in the RM *UBR1* promoter that affects the degradation of Type 2 Arg/N-end degrons, we engineered BY strains harboring the two individual RM promoter variants, -469A>T and -197T>G. We tested the influence of these variants on the degradation rate of the Phe and Trp N-degrons. The -469A>T variant significantly increased the degradation rate of both the Phe and Trp N-degrons (Figure 5C, Supplementary Table 5). The -197T>G did not alter the degradation of either N-degron (Figure 5C, Supplementary Table 5). The magnitude of the effect caused by the -469A>T variant suggests that this variant accounts for the majority of the *UBR1* QTL's effects on the degradation of Type 2 Arg/N-end substrates (Figure 5B/C).

Discussion

The activity of the ubiquitin-proteasome system (UPS) is dynamic and shaped by a diverse array of regulatory mechanisms[1,3,33]. Yet, to what extent genetic variation shapes UPS activity and associated regulatory mechanisms is almost entirely unknown. We combined recent advances in the design of fluorescent reporters of UPS protein degradation[48,49] with a powerful genetic mapping method[44,71,78] to systematically characterize genetic influences on UPS activity in the yeast *Saccharomyces cerevisiae*.

Our results establish several principles for understanding how individual genetic differences create heritable variation in UPS activity. Most prominently, UPS activity is a genetically complex trait, shaped by variation at many loci throughout the genome. We identified 149 UPS activity QTLs across 20 N-degrons, with a median of 7 QTLs per degron (Figure 4B, Supplementary Table 2). Similar to many other genetically complex traits[40–42], the distribution of QTL effect sizes was continuous and composed of many loci with small effects and few loci with large effects. At a majority of QTLs, the RM allele was associated with higher UPS activity. This excess matches our observation that RM had significantly higher UPS activity for most N-degrons in both the Ac/N-end and Arg/N-end pathways (Figure 2D, Supplementary Figures 2 and 3). Collectively, these results suggest that we have identified key sources of the genetic basis of the difference in UPS activity in these two strains and that heritable differences in UPS activity can arise through multiple, distinct genetic architectures. This considerable complexity in yeast suggests that a similar diversity of complex architectures may shape the activity of human UPS pathways. In humans, rare individual missense mutations that ablate the function of UPS genes cause a variety of diseases, including neurodegenerative, immune, and metabolic disorders, by impairing the normal regulation of UPS activity[14,21,23,28]. Our results are consistent with the notion that the collective effects of many loci that more subtly perturb the regulation of UPS activity may contribute to the genetic risk for non-Mendelian forms of these diseases[79].

Many of the QTLs we discovered were specific to the Ac/N-end or Arg/N-end pathways. Substrates for these pathways are recognized and processed through unique sets of molecular effectors[33,57,58]. Differences in the number and patterns of loci we detected between pathways likely reflect the distinct mechanisms by which Arg/N-end and Ac/N-end substrates are recognized and degraded and highlights the complexity of genetic influences of UPS activity.

Reporters for the Ac/N-end pathway had a higher number of QTLs per reporter than those of the Arg/N-end pathway. This difference was driven, in part, by a set of QTLs that influenced the majority or all of the Ac/N-end degrons. By contrast, we detected no QTLs that influenced all 12 Arg/N-degrons and only a few QTLs that influenced the majority of Arg/N-degrons. Unlike the shared Ac/N-end QTLs, many of the Arg/N-end QTLs that were detected with multiple reporters had divergent effects on Type 1 and Type 2 Arg/N-degrons. These results suggest that genetic effects on Ac/N-end substrates may be more general than those influencing Arg/N-end substrates.

Ac/N-degrons are generated by the co-translational excision of N-terminal Met residues, followed by acetylation of the newly exposed N-terminal amino acid[33,50,58]. The Met amino peptidases and N-terminal acetyltransferases that mediate these reactions act broadly on Ac/N-degrons[33] and our results suggest that genetic influences on Ac/N-end pathway substrates are similarly broad in

their effects. Arg/N-degrons comprise the set of amino acids with side chains too large to accommodate Met amino peptidases[33,80]. Instead, Arg/N-degrons appear result from cleavage by proteases other than Met-aminopeptidases, though the mechanisms of this process are less well-characterized than those that produce Ac/N-degrons[33]. Some Arg/N-end degrons undergo further modification in a manner analogous to the N-terminal acetylation of Ac/N-degrons, but these processes affect only a handful of N-terminal amino acids. For example, N-terminal Asn and Gln residues are deamidated by Nta1p to produce N-terminal Asp and Glu residues, which are subsequently arginylated to yield functional Arg/N-degrons[81]. We identified a QTL centered on the *NTA1* gene for our Asn TFT, suggesting the processing of pro-Arg/N-degrons may also be shaped by degron- and pathway-specific genetic variation. More broadly, the high level of pathway-specificity in our set of N-end rule QTLs raises the possibility that other UPS pathways are also shaped by unique patterns of genetic influences on the sequence of molecular events comprising the recognition, processing, and degradation of their substrates.

We showed that an Arg/N-end-specific QTL was caused by variation in the *UBR1* gene, which encodes the E3 ligase of the Arg/N-end pathway. The QTL had divergent effects on Type 1 and Type 2 Ubr1p Arg/N-end substrates, wherein the RM allele increased the degradation of Type 2 substrates and the Type 1 His N-degron and decreased the degradation of the other Type 1 substrates (Figure 5B). Remarkably, *UBR1* harbored multiple causal variants in the *UBR1* promoter, ORF, and terminator that each differentially affected Type 1 and 2 Arg/N-degrons (Figure 5A/B). We identified the causal variant in the *UBR1* promoter as an A to T substitution at position -469 (Figure 5C). A previous study showed that the this variant increases gene expression as measured in a synthetic reporter system[82]. Taken together, these results suggest that the -469 variant alters Arg/N-end pathway activity by increasing the expression of *UBR1*. Moreover, the existence of at least two additional causal variants in the *UBR1* ORF and terminator demonstrates that genetic effects on UPS activity engage diverse regulatory mechanisms, even within a single gene.

Our results add to an emerging picture of how variation in *UBR1* shapes a variety of cellular and organismal traits. Mutations in human *UBR1* cause Johanson-Blizzard sydrome (JBS)[26], an autosomal recessive disorder associated with a variety of congenital abnormalities, including pancreatic insufficiency, cognitive defects, and morphological abnormalities[26,83]. The majority of JBS-causing mutations are nonsense, frameshift, and splice site variants that abolish Ubr1 function[79,83]. A series of missense mutations in *UBR1* also cause JBS. However, the associated phenotypes in these patients are much less severe and the mutant forms of *UBR1* still retain some activity towards their cognate degrons[79,84]. Our results suggest that these findings represent one extreme of a continuous distribution of variant effects on *UBR1*. A more complete understanding of the genetic regulation of *UBR1* activity could provide important insights into the many physiological processes regulated

by the Arg/N-end pathway[6,33]. We anticipate that the full complement of genetic effects on *UBR1* abundance and function in humans will be large and comprise diverse molecular mechanisms of effect.

Finally, the genetic mapping strategy we have developed provides a generalized framework for mapping genetic effects on protein degradation with quantitative precision and high statistical power. Extending our approach to additional protein degradation pathways may reveal additional pathway-specific genetic architectures. Similarly, our approach represents a viable strategy for mapping genetic effects on the degradation of individual proteins. Considerable discrepancies exist between genetic influences on gene expression at the mRNA and protein levels[42,71,73]. Many protein-specific QTLs could result from genetic effects on protein degradation, either through loci that alter the activity of protein degradation pathways or the degradation of individual proteins.

Our genetic mapping strategy revealed that UPS activity is a genetically complex trait, shaped by variation throughout the genome. Individual loci influence UPS activity through diverse molecular mechanisms and often in a pathway-specific manner. Our results lay the groundwork for future efforts to understand how heritable differences in UPS activity contribute to variation in complex cellular and organismal traits, including the many diseases marked by aberrant UPS activity.

Materials and Methods

Tandem Fluorescent Timer (TFT) Ubiquitin-Proteasome System Activity Reporters

We used tandem fluorescent timers (TFTs) to measure ubiquitin-proteasome system (UPS) activity. TFTs are fusions of two fluorescent proteins with distinct spectral profiles and maturation kinetics[48,49]. In the most common implementation, a TFT consists of a faster maturing green fluorescent protein (GFP) and a slower maturing red fluorescent protein (RFP). Because the FPs in the TFT mature at different rates, the RFP/GFP ratio (hereafter, "TFT ratio") changes over time. If the RFP's maturation rate exceeds the TFT's degradation rate, then the TFT ratio is proportional to the construct's degradation rate[48,49] (Figure 1). Thus, when comparing two TFT-tagged substrates, a lower TFT ratio reflects a faster degradation rate and a higher TFT ratio reflects a slower degradation rate.

We used fluorescent proteins from previously characterized TFTs in our experiments. superfolder GFP[85] (sfGFP) was used as the GFP in all TFTs. sfGFP matures in approximately 5 minutes and has excitation and emission maximums of 485 nm and 510 nm, respectively[85]. The

RFP in our TFTs was either mCherry or mRuby. mCherry matures in approximately 40 minutes and has excitation and emission maximums of 587 nm and 610 nm, respectively[86]. The mCherry-sfGFP TFT can detect degradation rate differences in substrates with lifetimes of approximately 80 minutes[49,54]. mRuby matures in approximately 170 minutes and has excitation and emission maximums of 558 nm and 605 nm, respectively[87]. The mRuby-sfGFP TFT can detect degradation rate differences in substrates with lifetimes of approximately 340 minutes, although it is less sensitive than the mCherry-sfGFP for substrates with half-lives less than 80 minutes[49,54]. All TFT fluorescent proteins are monomeric. We separated green and red FPs in each TFT with an unstructured 35 amino acid linker sequence to minimize fluorescence resonance energy transfer (FRET)[49].

Construction of Arg/N-end and Ac/N-end Pathway TFTs

We built and characterized TFTs capable of measuring the activity of the UPS N-end rule pathway[33,56] which relates a protein's degradation rate to the identity of its N-degron[55]). To generate TFT constructs with defined N-terminal amino acids, we used the ubiquitin-fusion technique[33,56,60], which involves placing a ubiquitin moiety immediately upstream of a sequence encoding the desired N-degron. During translation, ubiquitin-hydrolases cleave the ubiquitin moiety, exposing the N-degron. We obtained DNA encoding the *Saccharomyces cerevisiae* ubiquitin sequence and a peptide linker sequence derived from *Escherichia coli* β-galactosidase previously used to identify components of the Arg/N-end and Ac/N-end pathways[56] by DNA synthesis (Integrated DNA Technologies [IDT], Coralville, Iowa, USA). The peptide linker sequence is unstructured and contains internal lysine residues required for ubiquitination and degradation by the UPS[56,58]. Peptide linkers encoding the 20 possible N-terminal amino acids were made by PCR amplifying the linker sequence using oligonucleotides encoding each unique N-terminal amino acid (Supplementary Table 6).

We then devised a general strategy to assemble TFT-containing plasmids with defined N-terminal amino acids (Supplementary Figure 1). We first obtained sequences encoding each reporter element by PCR or gene synthesis. We codon-optimized the sfGFP, mCherry, mRuby, FMDV2A, and the TFT linker sequences for expression in *S. cerevisiae* using the Java Codon Adaptation Tool (JCaT)[88] and purchased double-stranded synthetic DNA fragments of each sequence (IDT; Coralville, IA, USA). We used the *TDH3* promoter to drive expression of each TFT reporter. The *TDH3* promoter was PCR-amplified from Addgene plasmid #67639 (a gift from John Wyrick). We used the *ADH1* terminator for all TFT constructs, which we PCR amplified from Addgene plasmid #67639. We used the KanMX cassette[89] as the selection module for all TFT reporters and obtained this sequence by PCR amplification of Addgene plasmid #41030 (a gift from Michael Boddy). Thus, each construct has the general structure of *TDH3* promoter, N-degron, linker sequence, TFT module, *ADH1* terminator, and the KanMX resistance cassette (Supplementary

Figure 1).

We used Addgene plasmid #35121 (a gift from John McCusker) as the plasmid DNA backbone for all TFT constructs. Digesting this plasmid with BamHI and EcoRV restriction enzymes produces a 2451 bp fragment that we used as a vector backbone for TFT plasmid assembly. We obtained a DNA fragment containing 734 bp of sequence upstream of the *LYP1* start codon, a SwaI restriction site, and 380 bp of sequence downstream of the *LYP1* stop codon synthesis (IDT, Coralville, IA, USA). We performed isothermal assembly cloning using the New England Biolabs (NEB) HiFi Assembly Cloning Kit (NEB, Ipswhich, MA, USA) to insert the *LYP1* homology sequence into the BamHI/EcoRV digest of Addgene plasmid #35121 to create the final backbone plasmid BFA0190 (Supplementary Figure 1). We then combined SwaI digested BFA0190 and the components of each TFT reporter and used the NEB HiFi Assembly Kit (NEB; Ipswhich, MA, USA) to produce each TFT plasmid. The 5' and 3' *LYP1* sequences in each TFT contain naturally-occurring SacI and BglII restriction sites, respectively. We digested each TFT with SacI and BglII (NEB; Ipswhich, MA, USA) to obtain a linear DNA transformation fragment. The flanking *LYP1* homology and kanMX module in each TFT construct allows selection for reporter integration at the *LYP1* locus using G418[90] and the toxic amino acid analogue thialysine (S-(2-aminoethyl)-L-cysteine hydrochloride)[66,67,91](Supplementary Figure 1). The sequence identity of all assembled plasmids was verified by Sanger sequencing. The full list of plasmids used in this study is found in Supplementary Table 8.

Yeast Strains and Handling

Yeast Strains

We used two strains of the yeast *Saccharomyces cerevisiae* to characterize UPS activity reporters and perform genetic mapping of UPS activity. The BY strain is haploid with genotype MATa his3Δ hoΔ and is closely related to the *S. cerevisiae* S288C laboratory strain. The second strain, RM is a haploid strain with genotype MATα can1Δ::STE2pr-SpHIS5 his3Δ::NatMX AMN1-BY hoΔ::HphMX URA3-FY and was isolated from a California vineyard. BY and RM differ at 1 nucleotide per 200 base pairs on average and approximately 45,000 single nucleotide variants (SNVs) between the strains can serve as markers in a genetic mapping experiment[43,44,71,73].

We built additional strains for characterizing our UPS activity reporters by deleting individual UPS genes from the BY strain. Each deletion strain was constructed by replacing the targeted gene with the NatMX cassette[90], which confers resistance to the antibiotic nourseothricin. We PCR amplified the NatMX cassette using Addgene plasmid #35121 with primers with homology to the 5' upstream and 3' downstream sequences of the targeted gene. The oligonucleotides for

Short Name	Genotype	Antibiotic Resistance	Auxotrophies
BY	*MATa his3Δ hoΔ*		histidine
RM	*MATα can1Δ::STE2pr-SpHIS5 his3Δ::NatMX hoΔ::HphMX*	clonNAT, hygromycin	histidine
BY *rpn4Δ*	*MATa his3Δ hoΔ rpn4Δ::NatMX*	clonNAT	histidine
BY *ubr1Δ*	*MATa his3Δ hoΔ ubr1Δ::NatMX*	clonNAT	histidine
BY *doa10Δ*	*MATa his3Δ hoΔ doa10Δ::NatMX*	clonNAT	histidine

Table 1: Base strain genotypes

each gene deletion cassette amplification are listed in Supplementary Table 6. We created a BY strain lacking the *UBR1* gene, which encodes the Arg/N-end pathway E3 ligase Ubr1p. We refer to this strain hereafter as "BY ubr1Δ". We created a BY strain ("BY doa10Δ") lacking the *DOA10* gene that encodes the Ac/N-end pathway E3 ligase Doa10p. Finally, we created a BY strain ("BY rpn4Δ") lacking the *RPN4* that encodes the proteasome transcription factor Rpn4p. Table 1 lists these strains and their full genotypes. Supplementary Table 8 contains the complete list of strains used in this study.

Table 2 describes the media formulations used for all experiments. Synthetic complete amino acid powders (SC -lys and SC -his -lys -ura) were obtained from Sunrise Science (Knoxville, TN, USA). Where indicated, we added the following reagents at the indicated concentrations to yeast media: G418, 200 mg/mL (Fisher Scientific, Pittsburgh, PA, USA); clonNAT (nourseothricin sulfate, Fisher Scientific), 50 mg/L; thialysine (S-(2-aminoethyl)-L-cysteine hydrochloride; MillporeSigma, St. Louis, MO, USA), 50 mg/L; canavanine (L-canavanine sulfate, MillporeSigma), 50 mg/L.

Yeast Transformation

We used a standard yeast transformation protocol to construct reporter control strains and build strains with TFTs[92]. In brief, we inoculated yeast strains growing on solid YPD medium into 5 mL of YPD liquid medium for overnight growth at 30 °C. The following morning, we diluted 1 mL of saturated culture into 50 mL of fresh YPD and grew the cells for 4 hours. The cells were then successively washed in sterile ultrapure water and transformation solution 1 (10 mM Tris HCl [pH 8.0], 1 mM EDTA [pH 8.0], and 0.1 M lithium acetate). At each step, we pelleted the cells by centrifugation at 3000 rpm for 2 minutes in a benchtop centrifuge and discarded the supernatant. The cells were suspended in 100 μL of transformation solution 1 along with 50 μg

Media Name	Abbreviation	Formulation
Yeast-Peptone-Dextrose	YPD	10 g/L yeast extract
		20 g/L peptone
		20 g/L dextrose
Synthetic Complete	SC	6.7 g/L yeast nitrogen base
		1.96 g/L amino acid mix -lys
		20 g/L dextrose
Haploid Selection	SGA	6.7 g/L yeast nitrogen base
		1.74 g/L amino acid mix -his -lys -ura
		20 g/L dextrose
Sporulation	SPO	1 g/L yeast extract
		10 g/L potassium acetate
		0.5 g/L dextrose

Table 2: Media Formulations

of salmon sperm carrier DNA and 300 ng of transforming DNA. The cells were incubated at 30 °C for 30 minutes and 700 μL of transformation solution 2 (10 mM Tris HCl [pH 8.0], 1 mM EDTA [pH 8.0], and 0.1 M lithium acetate in 40% polyethylene glycol [PEG]) was added to each tube, followed by a 30 minute heat shock at 42 °C. We then washed the transformed cells in sterile, ultrapure water. We added 1 mL of liquid YPD medium to each tube and incubated the tubes for 90 minutes with rolling at 30 °C to allow for expression of the antibiotic resistance cassettes. After washing with sterile, ultrapure water, we plated 200 μL of cells on solid SC -lys medium with G418 and thialysine, and, for strains with the NatMX cassette, clonNAT. For each strain, we streaked 8 independent colonies (biological replicates) from the transformation plate for further analysis. We verified reporter integration and integration site by colony PCR[93]. The primers used for these experiments are listed in Supplementary Table 6.

Yeast Mating and Segregant Populations

We created populations of genetically variable, recombinant cells ("segregants") for genetic mapping using a modified synthetic genetic array (SGA) approach[66,67]. We first mated BY strains with a given TFT to RM by mixing freshly streaked cells of each strain on solid YPD medium. For each TFT, we mated two independently-derived clones (biological replicates) to the RM strain. Cells were grown overnight at 30 °C and we selected for diploid cells (successful BY-RM matings) by streaking mated cells onto solid YPD medium with G418 (which selects for the KanMX cassette in the TFT in the BY strain) and clonNAT (which selects for the NatMX cassette in the RM strain).

We inoculated 5 mL of YPD with freshly streaked diploid cells for overnight growth at 30 °C. The next day, we pelleted the cultures, washed them with sterile, ultrapure water, and resuspended the cells in 5 mL of SPO liquid medium (Table 2). We sporulated the cells by incubating them at room temperature with rolling for 9 days. After confirming sporulation by brightfield microscopy, we pelleted 2 mL of culture, washed cells with 1 mL of sterile, ultrapure water, and resuspended cells in 300 μL of 1 M sorbitol containing 3 U of Zymolyase lytic enzyme (United States Biological, Salem, MA, USA) to degrade ascal walls. Digestions were carried out at 30 °C with rolling for 2 hours. We then washed the spores with 1 mL of 1 M sorbitol, vortexed for 1 minute at the highest intensity setting, resuspended the cells in sterile ultrapure water, and confirmed the release of cells from ascii by brightfield microscopy. We plated 300 μl of cells onto solid SGA medium containing G418 and canavanine. This media formulation selects for haploid cells with (1) a TFT via G418, (2) the *MATa* mating type via the *Schizosaccharomyces pombe HIS5* gene under the control of the *STE2* promoter (which is only active in *MATa* cells), and (3) replacement of the *CAN1* gene with *S. pombe HIS5* via the toxic arginine analog canavanine[66,67]. Haploid segregant populations were grown for 2 days at 30 °C and harvested by adding 10 mL of sterile, ultrapure water and scraping the cells from each plate. We pelleted each cell suspension by centrifugation at 3000 rpm for 10 minutes and resuspended the cells in 1 mL of SGA medium. We added 450 μL of 40% (v/v) sterile glycerol solution to 750 μL of segregant culture and stored samples in screw cap cryovials at -80 30 °C. We stored 2 independent sporulations of each reporter (derived from our initial matings) as independent biological replicates.

Flow Cytometry and Fluorescence-Activated Cell Sorting

Flow Cytometry

Yeast strains were manually inoculated into 400 μl of liquid SC -lys medium with G418 and grown overnight in 2 mL 96 well plates at 30.0 °C with mixing using a MixMate (Eppendorf, Hamburg, Germany). The following morning, we inoculated a fresh 200 μL of G418-containing SC -lys media with 4 μL of saturated cultures from the overnight growth. Cells were grown for an additional 3 hours prior to analysis by flow cytometry. All flow cytometry experiments were performed on an LSR II flow cytometer (BD, Franklin Lakes, NJ, USA) equipped with a 20 mW 488 nm laser with 488/10 and 525/50 filters for measuring forward/side scatter and sfGFP, respectively, as well as a 40 mW 561 nm laser and a 610/20 filter for measuring mCherry and mRuby. Table 3 lists the parameters and settings that were used for all flow cytometry and fluorescence-activated cell sorting (FACS) experiments. We recorded 10,000 cells each from 8 biological replicates per clonal strain for our analyses of BY, RM, and reporter control strains.

We analyzed flow cytometry data using R (R Foundation for Statistical Computing, Vienna

Parameter	Laser Line (nm)	Laser Setting (V)	Filter
forward scatter (FSC)	488	500	488/10
side scatter (SSC)	488	275	488/10
sfGFP	488	500	525/50
mCherry	561	615	610/20
mRuby	561	615	610/20

Table 3: Flow cytometry and FACS settings.

Austria) and the flowCore R package[94]. We first filtered each flow cytometry dataset to include only those cells within 10% ± the forward scatter (a proxy for cell size) median. We empirically determined that this gating approach captured the central peak of cells in the FSC histogram. This gating approach also removed cellular debris, aggregates of multiple cells, and restricted our analyses to cells of the same approximate size.

To better characterize differences in the degradation rate of N-end rule substrates within and between our strains, we transformed our flow cytometry data as follows. We first scaled the $\log_2$ TFT ratio relative to the sample with lowest degradation rate. Following this transformation, the strain with lowest degradation rate (typically the E3 ligase deletion strain) has a degradation rate of approximately 0 and the now-scaled TFT ratio is directly proportional to the construct's degradation rate. To compare degradation rates between strains and individual TFTs, we then converted scaled TFT ratios to Z scores.

For our initial characterization of TFTs, we then extracted the mean of the remaining cells from each of 8 biological replicates and used these values for inferential statistics. We used a one-way ANOVA with Tukey's HSD post-hoc test to analyze all parameters and strains. For analyzing segregant populations obtained by FACS, we used the entire FSC-filtered populations of cells for inferential statistics and used ANOVA with Tukey's HSD post-hoc test to compare all populations.

Fluorescence-Activated Cell Sorting

We collected populations of segregant cells for QTL mapping using a previously-described approach for isolating phenotypically extreme cell populations by FACS[71,73]. Segregant populations were thawed approximately 16 hours prior to cell sorting and grown overnight in 5 mL of SGA medium containing G418 and canavanine. The following morning, 1 mL of cells from each segregant population was diluted into a fresh 4 mL of SGA medium containing G418 and canavanine. Segregant populations were then grown for an additional 4 hours prior to sorting. All FACS ex-

periments were carried out using a FACSAria II cell sorter (BD). We used plots of side scatter (SSC) height by side scatter height and forward scatter (FSC) height by forward scatter width to remove doublets from each sample. We then filtered cells on the basis of FSC area, restricting our sorts to ± 7.5% of the central FSC peak, which we empirically determined excluded cellular debris and aggregates and encompassed the primary haploid cell population. Finally, we defined a fluorescence-positive population by comparing each segregant population to negative control BY and RM strains without TFTs. We collected 20,000 cells each from 3 populations of cells for each segregant population:

1. Fluorescence-positive cells without further gating, which were used as unsorted, "null" populations
2. The 2% lower tail of the TFT distribution
3. The 2% upper tail of the TFT distribution

We collected independent biological replicates of each population for each reporter. Each population of 20,000 cells was collected into sterile 1.5 mL polypropylene tubes containing 1 mL of SGA medium and grown overnight at 30 °C with rolling. The next day, we mixed 750 μL of cells with 450 μL of 40% (v/v) glycerol and stored this mixture in 2 mL 96 well plates at −80 °C.

Genomic DNA Isolation and Whole-Genome Sequencing

Genomic DNA Isolation and Library Preparation

We extracted genomic DNA from sorted segregant populations for whole-genome sequencing. Deep-well plates containing glycerol stocks of segregant populations were thawed and 800 μL of each sample was pelleted by centrifugation at 3700 rpm for 10 minutes. We discarded the supernantant and resuspended cell pellets in 800 μL of a 1 M sorbitol solution containing 0.1 M EDTA, 14.3 mM β-mercaptoethanol, and 500 U of Zymolyase lytic enzyme to digest cell walls prior to DNA extraction. The digestion reaction was carried out by resuspending cell pellets with mixing at 1000 rpm for 2 minutes followed by incubation for 2 hours at 37 °C. When the digestion reaction finished, we discarded the supernatant, resuspended cells in 50 μL of phosphate buffered saline, and used the Quick-DNA 96 Plus kit (Zymo Research, Irvine, CA, USA) to extract genomic DNA. We followed the manufacturer's protocol to extract genomic DNA with the following modifications. We incubated cells in a 20 mg/mL proteinase K solution overnight with incubation at 55 °C. After completing the DNA extraction protocol, we eluted DNA using 40 μL of DNA elution buffer (10 mM Tris-HCl [pH 8.5], 0.1 mM EDTA). The DNA concentration for each sample was determined using the Qubit dsDNA BR assay kit (Thermo Fisher Scientific, Waltham, MA, USA) in a 96 well format using a Synergy H1 plate reader (BioTek Instruments, Winooski, VT, USA).

We used a previously-described approach to prepare libraries for short-read whole-genome sequencing on the Illumina Next-Seq platform[71,73]. We used the Nextera DNA library kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions with the following modifications. For the tagmentation reaction, 5 ng of genomic DNA from each sample was diluted in a master mix containing 4 μL of Tagment DNA buffer, 1 μL of sterile, molecular biology grade water, and 5 μL of Tagment DNA enzyme diluted 1:20 in Tagment DNA buffer. The tagmentation reaction was run on a SimpliAmp thermal cycler (Thermo Fisher Scientific) using the following parameters: 55 °C temperature, 20 μL reaction volume, 10 minute incubation. To prepare libraries for sequencing, we added 10 μL of the tagmentation reaction to a master mix containing 1 μL of an Illumina i5 and i7 index primer pair mixture, 0.375 μL of ExTaq polymerase (Takara Bio, Mountain View, CA, USA), 5 μL of ExTaq buffer, 4 μL of a dNTP mixture, and 29.625 μL of sterile molecular biology grade water. We generated all 96 possible index oligo combinations using 8 i5 and 12 i7 index primers. The library amplification reaction was run on a SimpliAmp thermal cycler with the following parameters: initial denaturation at 95 °C for 30 seconds, then 17 cycles of 95 °C for 10 seconds (denaturation), 62 °C for 30 seconds (annealing), and 72 °C for 3 minutes (extension). We quantified the DNA concentration of each reaction using the Qubit dsDNA BR assay kit (Thermo Fisher Scientific) and pooled 10 μL of each reaction. This pooled mixture was run on a 2% agarose gel and we extracted and purified DNA in the 400 bp to 600 bp region using the Monarch Gel Extraction Kit (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's instructions.

Whole-Genome Sequencing

We submitted pooled, purified DNA libraries to the University of Minnesota Genomics Center (UMGC) for Illumina sequencing. Prior to sequencing, UMGC staff performed three quality control assays. Library concentration was determined using the PicoGreen dsDNA quantification reagent (Thermo Fisher Scientific). Library size was determined using the Tapestation electrophoresis system (Agilent Technologies, Santa Clara, CA, USA) with libraries in the range of 200 to 700 bp passing QC. Library functionality was determined using the KAPA DNA Library Quantification kit (Roche, Penzberg, Germany), with libraries with a concentration greater than 2 nM passing. All submitted libraries passed each QC assay. We submitted 7 libraries for sequencing at different times. Libraries were sequenced on a NextSeq 550 instrument (Illumina). Depending on the number of samples, we used the following output settings. For libraries with 70 or more samples (2 libraries), 75 bp paired end sequencing was performed in high-output mode to generate approximately 360 $\times$ 10^6 reads. For libraries with 50 or fewer samples (5 libraries), 75 bp paired end sequencing was performed in mid-output mode to generate approximately 120 $\times$ 10^6 reads. Read coverage ranged from 9 to 35 with a median coverage of 28 across all libraries. Sequence data de-multiplexing was

performed by UMGC. Data are currently being deposited into the NIH Sequence Read Archive.

QTL Mapping

Raw Sequence Data Processing

We processed sequencing data to identify QTLs using a previously-described approach for genetic mapping by bulk segregant analysis[44,71,73]. We initially filtered reads to include only those reads with mapping quality scores greater than 30. We aligned the filtered reads to the *S. cerevisiae* reference genome (version sacCer3) using BWA[95] (command: "mem -t 24"). We then used samtools to remove mismatches and PCR duplicates (command: "samtools rmdup -S"). Finally, we produced vcf files containing coverage and allelic read counts at each of 18,871 high-confidence, reliable SNPs[44,96] (command: "samtools mpileup -vu -t INFO/AD -l"). Because the BY strain is closely related to the S288C *S. cerevisiae* strain, we used BY alleles as reference and RM alleles as alternative alleles.

QTL Mapping

We used the vcf files generated by our raw sequence read processing to detect UPS activity QTLs. We used the MULTIPOOL algorithm to identify significant QTLs[97]. MULTIPOOL estimates logarithm of the odds (LOD) scores, which we used to identify QTLs exceeding an empirically-derived significance threshold (see below). We used MULTIPOOL with the following settings: bp per centiMorgan = 2,200, bin size = 100 bp, effective pool size = 1,000. As in previous QTL mapping in the BY/RM cross by bulk segregant analysis, we excluded variants with alleles with frequencies higher than 0.9 or lower than 0.1[71,73]. We also used MULTIPOOL to estimate confidence intervals for each significant QTL, which we defined as a 2-LOD drop from the QTL peak position. To visualize QTLs and gauge their effects, we also computed the allele frequency differences (ΔAF) at each site between our high and low UPS activity pools. Because allele frequencies are affected by random counting noise, we used loess regression to smooth the allele frequency for each sample before computing ΔAF. We used the smoothed values to plot the ΔAF distribution and visualize the association of alleles with UPS activity.

Null Sorts and Empirical False Discovery Rate Estimation

We used a subset of our segregant populations to empirically estimate the false discovery rate (FDR) of our QTL mapping method. We collected 2 separate populations of 20,000 fluorescence-positive, unsorted cells from 8 independently-derived segregant populations. We included these populations in our whole-genome sequencing and used the resultant data to estimate the FDR. Because these cells are obtained from the same unsorted population in the same sample, any ΔAF differences between them are likely the result of technical noise or random variation. We permuted

these comparisons across segregant populations with the same TFT reporter for a total of 112 null comparisons. We define the "null QTL rate" at a given LOD threshold as the number of QTLs that exceeded the threshold in these comparisons divided by the number of null comparisons. To determine the FDR for a given LOD score, we then determined the number of QTLs for our experimental comparisons (extreme high TFT ratio versus extreme low TFT ratio). We define the "experimental QTL rate" as the number of experimental QTLs divided by the number of experimental comparisons. We then computed the FDR as follows:

$$null\ QTL\ rate = \frac{n.\ null\ QTLs}{n.\ null\ comparisons}$$

$$experimental\ QTL\ rate = \frac{n.\ experimental\ QTLs}{n.\ experimental\ comparisons}$$

$$FDR = \frac{null\ QTL\ rate}{experimental\ QTL\ rate}$$

We evaluated the FDR over a LOD range of 2.5 to 10 in 0.5 LOD increments. We found that a LOD value of 4.5 led to a null QTL rate of 0.0625 and an FDR of 0.507% and we used this value as our significance threshold for QTL mapping. We further filtered our QTL list by excluding QTLs that were not detected in each of two independent biological replicates. Replicating QTLs were defined as those whose peaks were within 100 kB of each other on the same chromosome with the same direction (positive or negative) of allele frequency difference between high and low UPS activity pools.

QTL Fine-Mapping

We used "CRISPR-Swap", a two-step method for scarless allelic editing, to fine-map QTLs to the level of their causal genes and nucleotides[77]. In the first step of CRISPR-Swap, a gene of interest (GOI) is deleted and replaced with a selectable marker. In the second step, cells are co-transformed with (1) a plasmid with CRISPR-cas9 and a guide RNA targeting the selectable marker used to remove the GOI and (2) a repair template encoding the desired allele of the GOI.

We used CRISPR-Swap to generate BY strains harboring the RM *UBR1* allele, as well as a series of chimeric BY/RM *UBR1* alleles, as described below. To do so, we first replaced the *UBR1* gene in BY with the NatMX selectable marker by transforming a PCR product encoding the NatMX cassette with 40 bp 5' and 3' overhangs homologous to *UBR1*. To generate the *NatMX::ubr1*Δ transformation fragment, we PCR amplified NatMX from Addgene plasmid #35121 using primers OFA1102 and OFA1103 (sequences in Supplementary Table 1) using Phusion Hot Start Flex DNA polymerase (NEB). The NatMX cassette was transformed into the BY strain using the methods

described above and transformants were plated onto YPD medium containing clonNAT. We verified the deletion of the wild-type *UBR1* allele from single-colony purified transformants by colony PCR (primer sequences listed in Supplementary Table 6).

We then modified the original CRISPR-Swap plasmid (PFA0055, Addgene plasmid #131774) to replace its *LEU2* selectable marker with the *HIS3* selectable marker, creating plasmid PFA0227 (Supplementary Table 7). To build PFA0277, we first digested PFA0055 with restriction enzymes BsmBI-v2 and HpaI to remove the *LEU2* selectable marker. We synthesized the *S. cerevisiae HIS3* selectable marker from plasmid pRS313[98] with 20 base pairs of overlap to BsmBI-v2/HpaI-digested PFA0055 on both ends. We used this synthetic *HIS3* fragment and BsmBI-v2/HpaI-digested PFA0055 to create plasmid PFA0227 by isothermal assembly cloning using the HiFi Assembly Cloning Kit (NEB) according to the manufacturer's instructions. In addition to the *HIS3* selectable marker, PFA0227 contains the cas9 gene driven by the constitutively active *TDH3* promoter and a guide RNA, gCASS5a, that directs cleavage of a site immediately upstream of the *TEF* promoter used to drive expression of the MX series of selectable markers[77,90]. We verified the sequence of PFA0227 by Sanger sequencing.

We used genomic DNA from BY and RM strains to use for PCR amplifying *UBR1* repair templates for the second step of CRISPR-Swap. Genomic DNA was extracted from BY and RM strains using the "10 minute prep" protocol[99]. We amplified full-length *UBR1* repair templates from RM and BY containing the gene's promoter, open-reading frame (ORF), and terminator using Phusion Hot Start Flex DNA polymerase (NEB). We also created chimeric repair templates containing combinations of BY and RM alleles using PCR splicing by overlap extension[100]. Table 4 lists the repair templates used for CRISPR swap:

Name	Promoter	ORF	Terminator
UBR1 BY	BY	BY	BY
UBR1 RM	RM	RM	RM
UBR1 RM pr	RM	BY	BY
UBR1 RM ORF	BY	RM	BY
UBR1 RM term	BY	BY	RM
UBR1 -469A>T	-469, RM; all other, BY	BY	BY
UBR1 -197T>G	-197, RM; all other, BY	BY	BY

Table 4: *UBR1* CRISPR-Swap repair templates.

The sequence of all repair templates was verified by Sanger sequencing.

To create *UBR1* allele swap strains, we co-transformed BY strains with 200 ng of plasmid PFA0227 and 1.5 μg of *UBR1* repair template. Transformants were selected and single colony purified on synthetic complete medium lacking histidine and then patched onto solid YPD medium. We tested each strain for the desired exchange of the NatMX selectable marker with a *UBR1* allele by patching strains onto solid YPD medium containing clonNAT. We then verified allelic exchange in strains lacking ClonNAT resistance by colony PCR. We kept 8 independently-derived biological replicates of each allele swap strain. To test the effects of each allele swap, we transformed a subset of TFTs into our allele swap strains and characterized TFT reporter activity by flow cytometry using the methods described above.

Data Analysis

All data were analyzed using R (version 3.6). Computational scripts used to process data, for statistical analysis and to generate figures are available at `http://www.github.com/mac230/UPS_QTL_paper`. Final figures and illustrations were made using Inkscape.

Acknowledgements

We thank Leonid Kruglyak for the BY and RM yeast strains and Michael Knop for technical assistance in implementing the TFT reporter system. We thank the University of Minnesota's Flow Cytometry Resource and Genomics Center for their contributions to the project. We thank the members of the Albert laboratory and the BioKansas Scientific Writing Program for critical feedback on the manuscript.

Competing Interests

The authors declare that they have no competing interests.

Funding

This work was supported by NIH grants F32-GM128302 to MAC and R35-GM124676 to FWA, as well as a Pew Scholarship in the Biomedical Sciences from the Pew Charitable Trusts to FWA.

Supplementary Materials

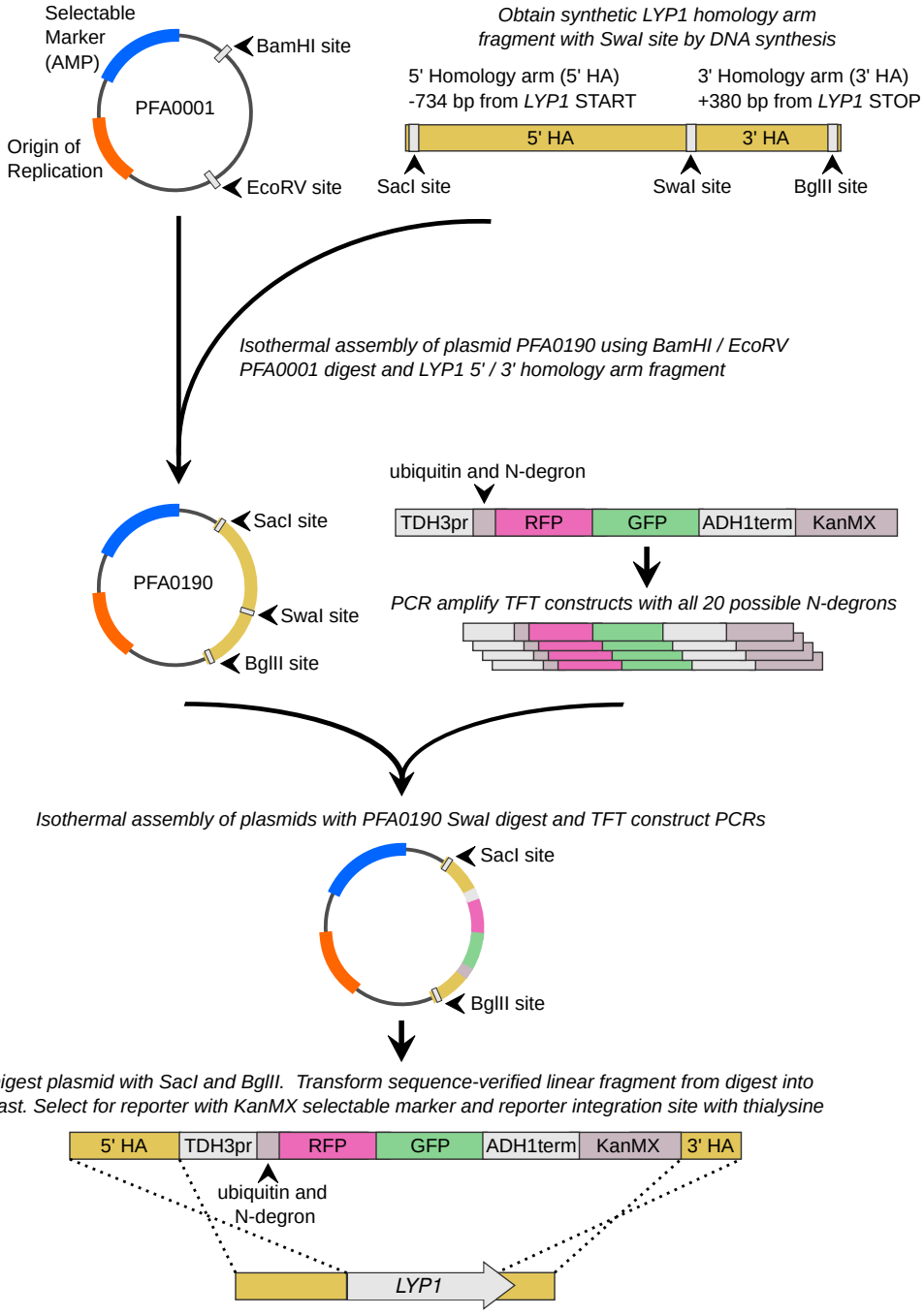
The following supplementary materials are included with this manuscript:

Supplementary Figures 1-6

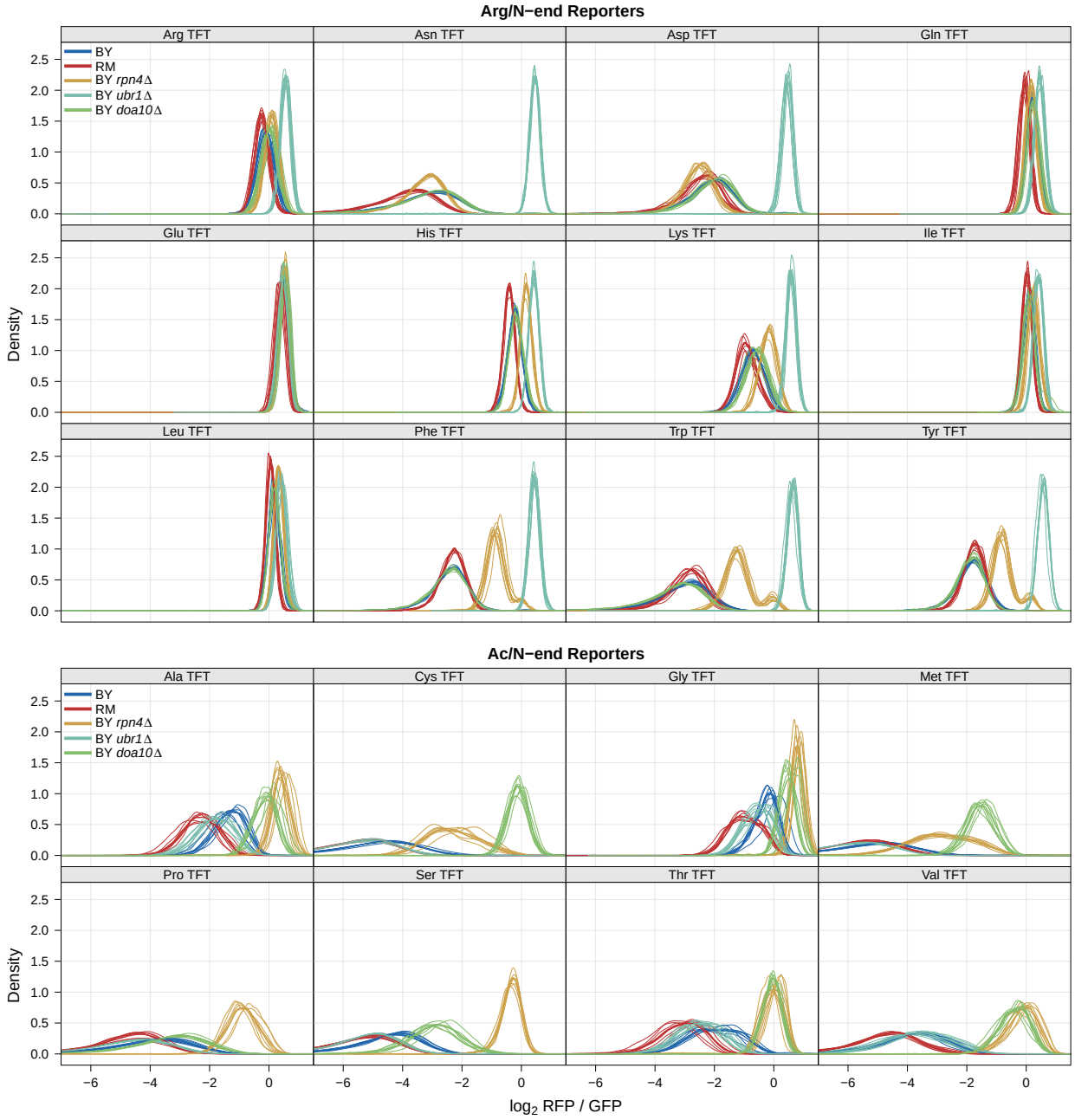
Supplementary Tables 1-8

Supplementary File 1: Allele frequency difference and LOD traces for the 20 N-degron TFT QTL mapping experiments

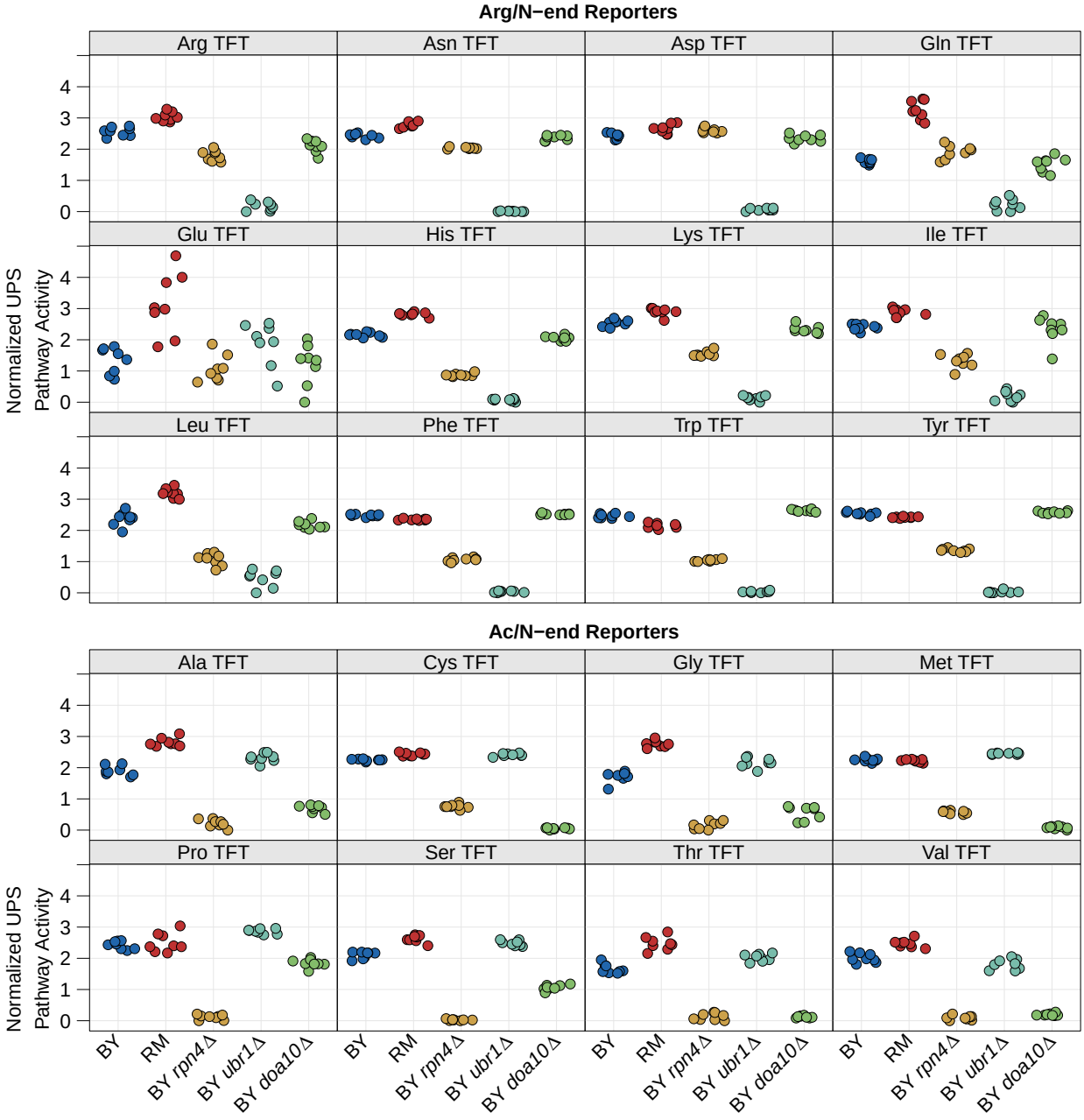
Supplementary Figures



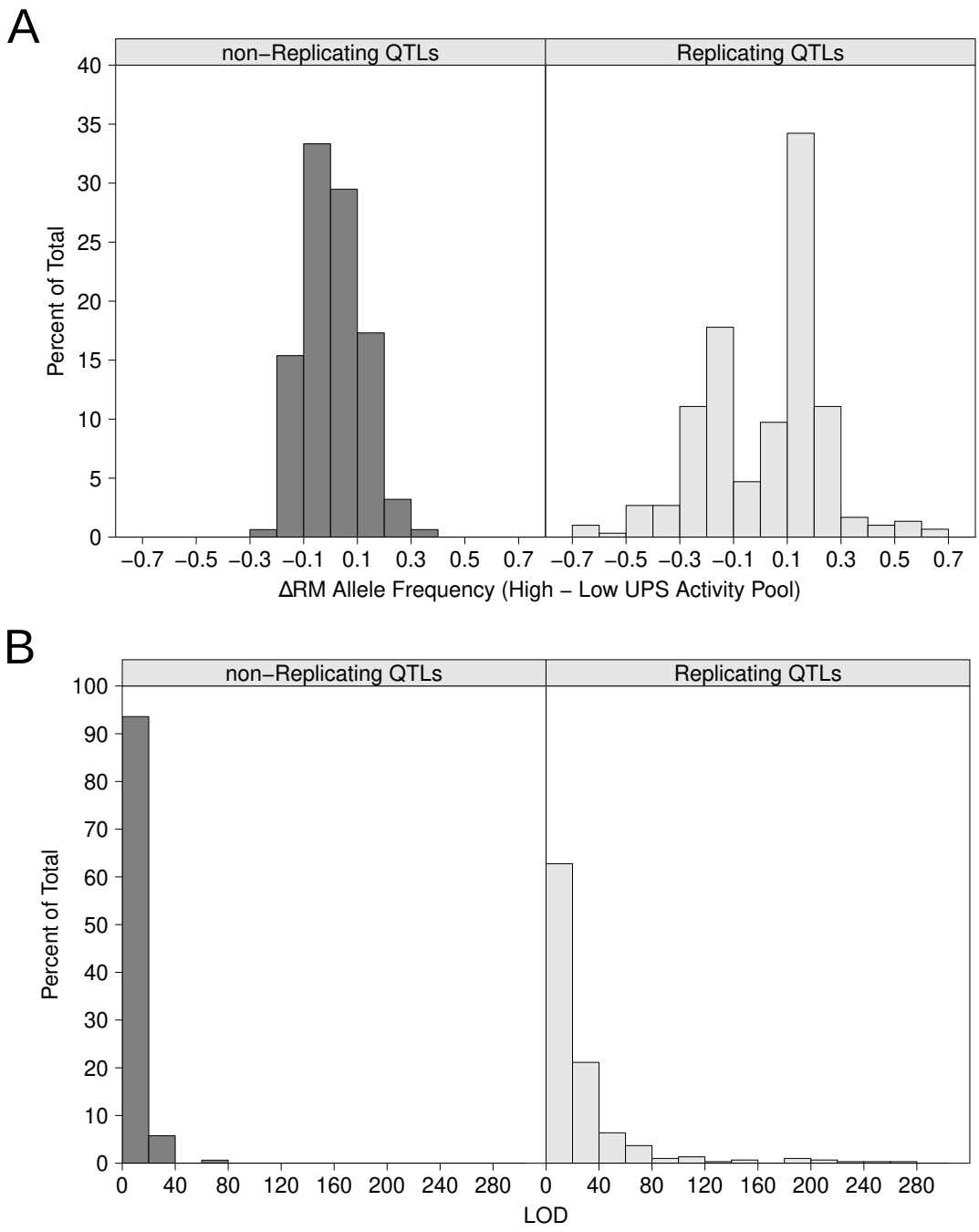
Supplementary Figure 1: Overview of the constructs and strain construction steps used to generate yeast strains harboring TFT UPS activity reporters.



Supplementary Figure 2: TFT Ratio Density Plots. TFT ratio plots for the full set of 20 N-degrons and strains are shown. N-degrons are separated by pathway. Because we used the same flow cytometry settings to acquire all data, there is some fluctuation around 0 for the RFP/GFP ratios for strains in which the TFT is stabilized. The thin lines show the density values for 10,000 cells each for each of 8 independent biological replicates per strain per reporter. Thick lines show the average for each strain and reporter.

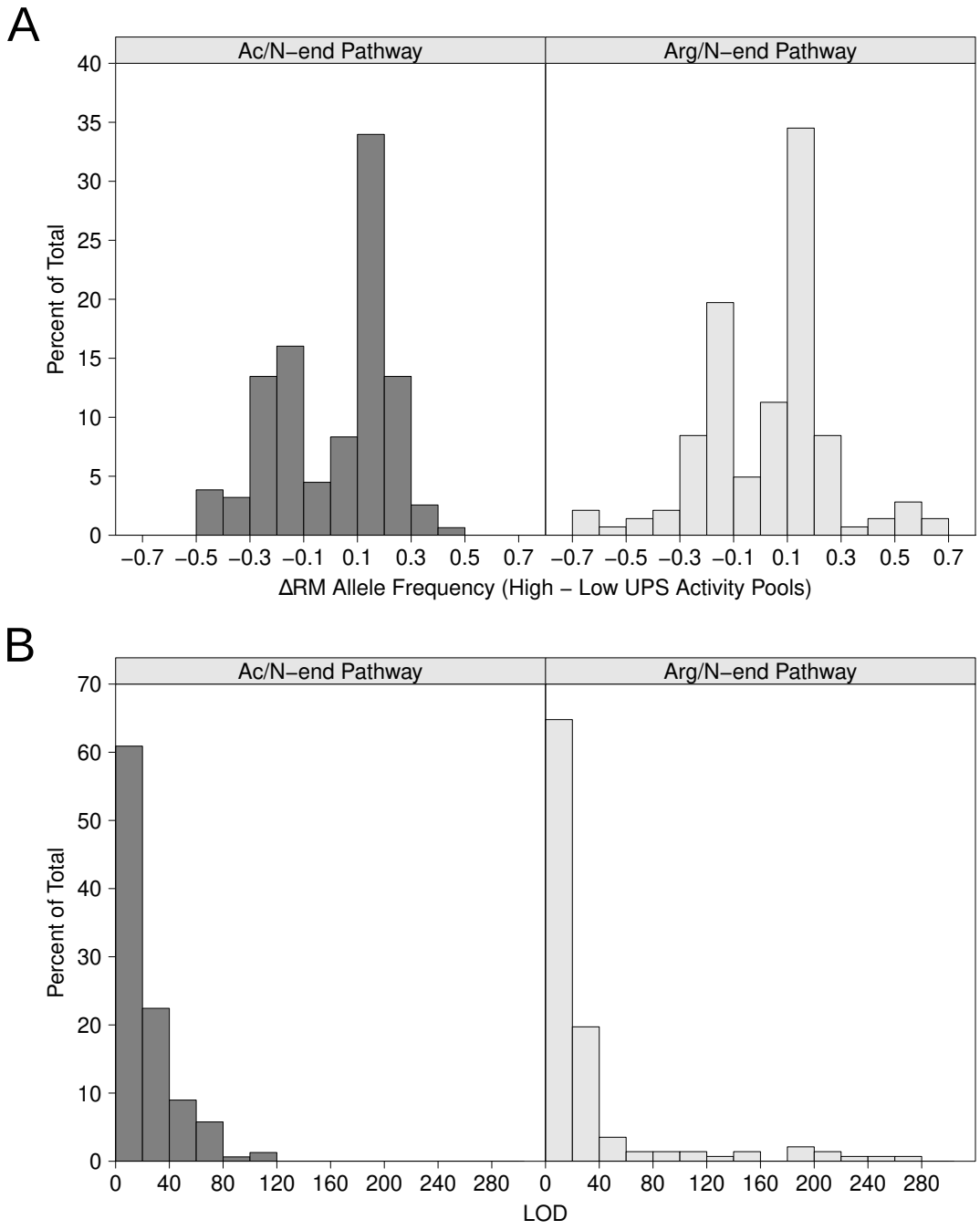


Supplementary Figure 3: N-Degron Strip Plots. The UPS activity for each N-degron and strain are shown as strip plots. N-degrons are separated by pathway. The median value was extracted from 10,000 cells from each of 8 independent biological replicates per strain per reporter. The data was converted to Z-scores and scaled such that high values correspond to high UPS activity and low values correspond to low UPS activity.

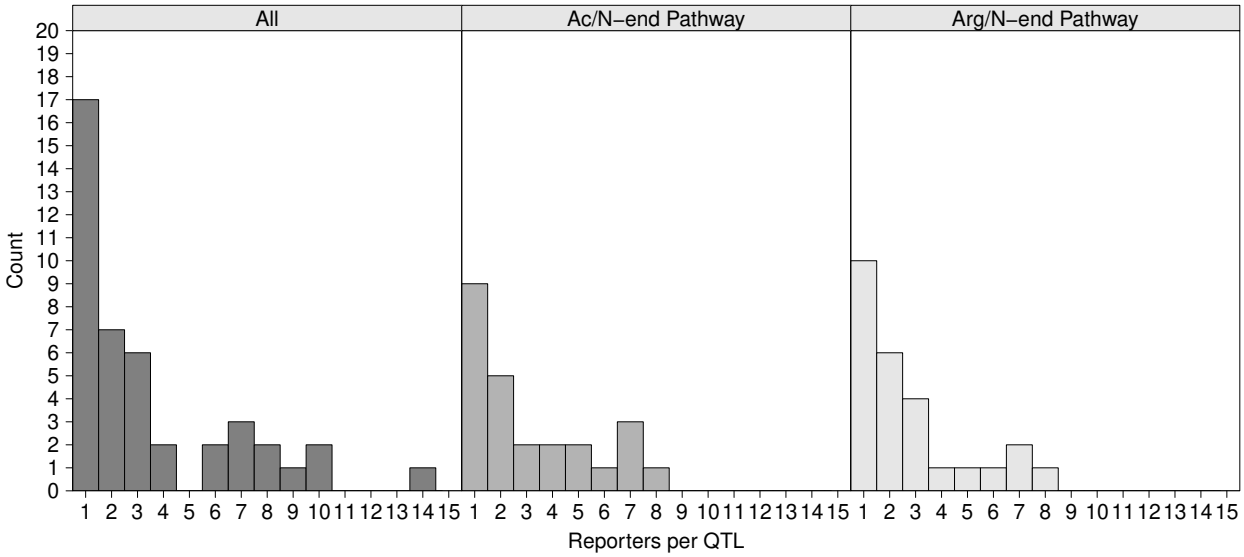


842

843 **Supplementary Figure 4:** Analysis of replicating and non-replicating UPS activity QTLs. A.
844 Allele frequency difference histograms for non-replicating (left) and replicating QTLs (right). B.
845 As in A, but for LOD score.



Supplementary Figure 5: Analysis of Ac/N-end and Arg/N-end QTLs. A. Allele frequency difference histograms for Ac/N-end (left) and Arg/N-end QTLs (right). B. As in A, but for LOD score.



Supplementary Figure 6: N-degron QTL Specificity. The histograms show the number of N-degrons affected by individual QTLs for all N-degrons (left), Ac/N-degrons (middle), and Arg/N-degrons (right).

References

[1] J. Hanna and D. Finley. A proteasome for all occasions. *FEBS Lett*, 581(15):2854–2861, Jun 2007.

[2] S. Bhattacharyya, H. Yu, C. Mim, and A. Matouschek. Regulated protein turnover: snapshots of the proteasome in action. *Nat Rev Mol Cell Biol*, 15(2):122–133, Feb 2014.

[3] C. Pohl and I. Dikic. Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science*, 366(6467):818–822, 11 2019.

[4] T. Tasaki and Y. T. Kwon. The mammalian N-end rule pathway: new insights into its components and physiological roles. *Trends Biochem Sci*, 32(11):520–528, Nov 2007.

[5] C. S. Hwang and A. Varshavsky. Regulation of peptide import through phosphorylation of Ubr1, the ubiquitin ligase of the N-end rule pathway. *Proc Natl Acad Sci U S A*, 105(49):19188–19193, Dec 2008.

[6] G. C. Turner, F. Du, and A. Varshavsky. Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway. *Nature*, 405(6786):579–583, Jun 2000.

[7] Z. Xia, G. C. Turner, C. S. Hwang, C. Byrd, and A. Varshavsky. Amino acids induce peptide uptake via accelerated degradation of CUP9, the transcriptional repressor of the PTR2 peptide transporter. *J Biol Chem*, 283(43):28958–28968, Oct 2008.

[8] R. G. Hu, H. Wang, Z. Xia, and A. Varshavsky. The N-end rule pathway is a sensor of heme. *Proc Natl Acad Sci U S A*, 105(1):76–81, Jan 2008.

[9] K. Madura and A. Varshavsky. Degradation of G alpha by the N-end rule pathway. *Science*, 265(5177):1454–1458, Sep 1994.

[10] N. Baloghova, T. Lidak, and L. Cermak. Ubiquitin Ligases Involved in the Regulation of Wnt, TGF-, and Notch Signaling Pathways and Their Roles in Mouse Development and Homeostasis. *Genes (Basel)*, 10(10), 10 2019.

[11] K. Flick and P. Kaiser. Protein degradation and the stress response. *Semin Cell Dev Biol*, 23(5):515–522, Jul 2012.

[12] S. Jain, J. R. Wheeler, R. W. Walters, A. Agrawal, A. Barsic, and R. Parker. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell*, 164(3):487–498, Jan 2016.

[13] J. S. Bett. Proteostasis regulation by the ubiquitin system. *Essays Biochem*, 60(2):143–151, 10 2016.

[14] B. Mészáros, M. Kumar, T. J. Gibson, B. Uyar, and Z. Dosztányi. Degrons in cancer. *Sci Signal*, 10(470), Mar 2017.

[15] C. W. Fhu and A. Ali. Dysregulation of the Ubiquitin Proteasome System in Human Malignancies: A Window for Therapeutic Intervention. *Cancers (Basel)*, 13(7), Mar 2021.

[16] N. P. Dantuma and L. C. Bott. The ubiquitin-proteasome system in neurodegenerative diseases: precipitating factor, yet part of the solution. *Front Mol Neurosci*, 7:70, 2014.

[17] J. M. Deger, J. E. Gerson, and R. Kayed. The interrelationship of proteasome impairment and oligomeric intermediates in neurodegeneration. *Aging Cell*, 14(5):715–724, Oct 2015.

[18] C. Zheng, T. Geetha, and J. R. Babu. Failure of ubiquitin proteasome system: risk for neurodegenerative diseases. *Neurodegener Dis*, 14(4):161–175, 2014.

[19] A. Nencioni, F. Grunebach, F. Patrone, A. Ballestrero, and P. Brossart. The proteasome and its inhibitors in immune regulation and immune disorders. *Crit Rev Immunol*, 26(6):487–498, 2006.

[20] A. L. Schwartz and A. Ciechanover. The ubiquitin-proteasome pathway and pathogenesis of human diseases. *Annu Rev Med*, 50:57–74, 1999.

[21] S. S. Wing. The UPS in diabetes and obesity. *BMC Biochem*, 9 Suppl 1:S6, Oct 2008.

[22] S. S. Thomas and W. E. Mitch. Mechanisms stimulating muscle wasting in chronic kidney disease: the roles of the ubiquitin-proteasome system and myostatin. *Clin Exp Nephrol*, 17(2):174–182, Apr 2013.

[23] D. Sepulveda-Falla, L. Chavez-Gutierrez, E. Portelius, J. I. Vélez, S. Dujardin, A. Barrera-Ocampo, F. Dinkel, C. Hagel, B. Puig, C. Mastronardi, F. Lopera, B. T. Hyman, K. Blennow, M. Arcos-Burgos, B. de Strooper, and M. Glatzel. A multifactorial model of pathology for age of onset heterogeneity in familial Alzheimer's disease. *Acta Neuropathol*, 141(2):217–233, 02 2021.

[24] T. Sjakste, N. Paramonova, I. Rumba-Rozenfelde, I. Trapina, O. Sugoka, and N. Sjakste. Juvenile idiopathic arthritis subtype- and sex-specific associations with genetic variants in the PSMA6/PSMC6/PSMA3 gene cluster. *Pediatr Neonatol*, 55(5):393–403, Oct 2014.

[25] I. Garcia-Martínez, C. Sánchez-Mora, M. Soler Artigas, P. Rovira, M. Pagerols, M. Corrales, E. Calvo-Sánchez, V. Richarte, M. Bustamante, J. Sunyer, B. Cormand, M. Casas, J. A. Ramos-Quiroga, and M. Ribasés. Gene-wide Association Study Reveals RNF122 Ubiquitin Ligase as a Novel Susceptibility Gene for Attention Deficit Hyperactivity Disorder. *Sci Rep*, 7(1):5407, 07 2017.

[26] M. Zenker, J. Mayerle, M. M. Lerch, A. Tagariello, K. Zerres, P. R. Durie, M. Beier, G. Hülskamp, C. Guzman, H. Rehder, F. A. Beemer, B. Hamel, P. Vanlieferinghen, R. Gershoni-Baruch, M. W. Vieira, M. Dumic, R. Auslender, V. L. Gil-da Silva-Lopes, S. Steinlicht, M. Rauh, S. A. Shalev, C. Thiel, A. B. Ekici, A. Winterpacht, Y. T. Kwon, A. Varshavsky, and A. Reis. Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). *Nat Genet*, 37(12):1345–1350, Dec 2005.

[27] K. Arima, A. Kinoshita, H. Mishima, N. Kanazawa, T. Kaneko, T. Mizushima, K. Ichinose, H. Nakamura, A. Tsujino, A. Kawakami, M. Matsunaka, S. Kasagi, S. Kawano, S. Kumagai, K. Ohmura, T. Mimori, M. Hirano, S. Ueno, K. Tanaka, M. Tanaka, I. Toyoshima, H. Sugino, A. Yamakawa, K. Tanaka, N. Niikawa, F. Furukawa, S. Murata, K. Eguchi, H. Ida, and K. Yoshiura. Proteasome assembly defect due to a proteasome subunit beta type 8 (PSMB8) mutation causes the autoinflammatory disorder, Nakajo-Nishimura syndrome. *Proc Natl Acad Sci U S A*, 108(36):14914–14919, Sep 2011.

[28] H. X. Deng, W. Chen, S. T. Hong, K. M. Boycott, G. H. Gorrie, N. Siddique, Y. Yang, F. Fecto, Y. Shi, H. Zhai, H. Jiang, M. Hirano, E. Rampersaud, G. H. Jansen, S. Donkervoort, E. H. Bigio, B. R. Brooks, K. Ajroud, R. L. Sufit, J. L. Haines, E. Mugnaini, M. A. Pericak-Vance, and T. Siddique. Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature*, 477(7363):211–215, Aug 2011.

[29] A. Varshavsky. Naming a targeting signal. *Cell*, 64(1):13–15, Jan 1991.

[30] A. Ciechanover, A. Orian, and A. L. Schwartz. Ubiquitin-mediated proteolysis: biological regulation via destruction. *Bioessays*, 22(5):442–451, May 2000.

[31] A. Hershko and A. Ciechanover. The ubiquitin system. *Annu Rev Biochem*, 67:425–479, 1998.

[32] J. A. M. Bard, E. A. Goodall, E. R. Greene, E. Jonsson, K. C. Dong, and A. Martin. Structure and Function of the 26S Proteasome. *Annu Rev Biochem*, 87:697–724, 06 2018.

[33] A. Varshavsky. The N-end rule pathway and regulation by proteolysis. *Protein Sci.*, 20(8):1298–1345, Aug 2011.

[34] G. Mannhaupt, R. Schnall, V. Karpov, I. Vetter, and H. Feldmann. Rpn4p acts as a transcription factor by binding to PACE, a nonamer box found upstream of 26S proteasomal and other genes in yeast. *FEBS Lett*, 450(1-2):27–34, Apr 1999.

[35] Y. Xie and A. Varshavsky. RPN4 is a ligand, substrate, and transcriptional regulator of the 26S proteasome: a negative feedback circuit. *Proc. Natl. Acad. Sci. U.S.A.*, 98(6):3056–3061, Mar 2001.

[36] J. J. S. VerPlank and A. L. Goldberg. Regulating protein breakdown through proteasome phosphorylation. *Biochem J*, 474(19):3355–3371, 09 2017.

[37] G. A. Collins and A. L. Goldberg. The Logic of the 26S Proteasome. *Cell*, 169(5):792–806, May 2017.

[38] K. Zientara-Rytter and A. Sirko. To deliver or to degrade - an interplay of the ubiquitin-proteasome system, autophagy and vesicular transport in plants. *FEBS J*, 283(19):3534–3555, 10 2016.

[39] R. S. Marshall and R. D. Vierstra. Dynamic Regulation of the 26S Proteasome: From Synthesis to Degradation. *Front Mol Biosci*, 6:40, 2019.

[40] T. F. Mackay, E. A. Stone, and J. F. Ayroles. The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet*, 10(8):565–577, Aug 2009.

[41] M. Civelek and A. J. Lusis. Systems genetics approaches to understand complex traits. *Nat Rev Genet*, 15(1):34–48, Jan 2014.

[42] F. W. Albert and L. Kruglyak. The role of regulatory variation in complex traits and disease. *Nat. Rev. Genet.*, 16(4):197–212, Apr 2015.

[43] I. M. Ehrenreich, J. P. Gerke, and L. Kruglyak. Genetic dissection of complex traits in yeast: insights from studies of gene expression and other phenotypes in the BYxRM cross. *Cold Spring Harb Symp Quant Biol*, 74:145–153, 2009.

[44] I. M. Ehrenreich, N. Torabi, Y. Jia, J. Kent, S. Martis, J. A. Shapiro, D. Gresham, A. A. Caudy, and L. Kruglyak. Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature*, 464(7291):1039–1042, Apr 2010.

[45] F. W. Albert, J. S. Bloom, J. Siegel, L. Day, and L. Kruglyak. Genetics of trans-regulatory variation in gene expression. *Elife*, 7, 07 2018.

[46] M. A. Eldeeb, R. Siva-Piragasam, M. A. Ragheb, M. Esmaili, M. Salla, and R. P. Fahlman. A molecular toolbox for studying protein degradation in mammalian cells. *J Neurochem*, 151(4):520–533, 2019.

[47] H. Ella, Y. Reiss, and T. Ravid. The Hunt for Degrons of the 26S Proteasome. *Biomolecules*, 9(6), 06 2019.

[48] A. Khmelinskii and M. Knop. Analysis of protein dynamics with tandem fluorescent protein timers. *Methods Mol. Biol.*, 1174:195–210, 2014.

[49] A. Khmelinskii, P. J. Keller, A. Bartosik, M. Meurer, J. D. Barry, B. R. Mardin, A. Kaufmann, S. Trautmann, M. Wachsmuth, G. Pereira, W. Huber, E. Schiebel, and M. Knop. Tandem fluorescent protein timers for in vivo analysis of protein dynamics. *Nat. Biotechnol.*, 30(7):708–714, Jun 2012.

[50] I. Kats, A. Khmelinskii, M. Kschonsak, F. Huber, R. A. Knieß, A. Bartosik, and M. Knop. Mapping Degradation Signals and Pathways in a Eukaryotic N-terminome. *Mol Cell*, 70(3):488–501, 05 2018.

[51] I. Koren, R. T. Timms, T. Kula, Q. Xu, M. Z. Li, and S. J. Elledge. The Eukaryotic Proteome Is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons. *Cell*, 173(7):1622–1635, 06 2018.

[52] H. C. Lin, C. W. Yeh, Y. F. Chen, T. T. Lee, P. Y. Hsieh, D. V. Rusnac, S. Y. Lin, S. J. Elledge, N. Zheng, and H. S. Yen. C-Terminal End-Directed Protein Elimination by CRL2 Ubiquitin Ligases. *Mol. Cell*, 70(4):602–613, 05 2018.

[53] A. Khmelinskii, E. Blaszczak, M. Pantazopoulou, B. Fischer, D. J. Omnus, G. Le Dez, A. Brossard, A. Gunnarsson, J. D. Barry, M. Meurer, D. Kirrmaier, C. Boone, W. Huber, G. Rabut, P. O. Ljungdahl, and M. Knop. Protein quality control at the inner nuclear membrane. *Nature*, 516(7531):410–413, Dec 2014.

[54] A. Khmelinskii, M. Meurer, C. T. Ho, B. Besenbeck, J. F?ller, M. K. Lemberg, B. Bukau, A. Mogk, and M. Knop. Incomplete proteasomal degradation of green fluorescent proteins in the context of tandem fluorescent protein timers. *Mol. Biol. Cell*, 27(2):360–370, Jan 2016.

[55] A. Varshavsky. N-degron and C-degron pathways of protein degradation. *Proc Natl Acad Sci U S A*, 116(2):358–366, 01 2019.

[56] A. Bachmair, D. Finley, and A. Varshavsky. In vivo half-life of a protein is a function of its amino-terminal residue. *Science*, 234(4773):179–186, Oct 1986.

[57] B. Bartel, I. W?nning, and A. Varshavsky. The recognition component of the N-end rule pathway. *EMBO J.*, 9(10):3179–3189, Oct 1990.

[58] C. S. Hwang, A. Shemorry, and A. Varshavsky. N-terminal acetylation of cellular proteins creates specific degradation signals. *Science*, 327(5968):973–977, Feb 2010.

[59] A. Bachmair and A. Varshavsky. The degradation signal in a short-lived protein. *Cell*, 56(6):1019–1032, Mar 1989.

[60] A. Varshavsky. Ubiquitin fusion technique and related methods. *Meth. Enzymol.*, 399:777–799, 2005.

[61] R. B. Brem and L. Kruglyak. The landscape of genetic complexity across 5,700 gene expression traits in yeast. *Proc Natl Acad Sci U S A*, 102(5):1572–1577, Feb 2005.

[62] S. Treusch, F. W. Albert, J. S. Bloom, I. E. Kotenko, and L. Kruglyak. Genetic mapping of MAPK-mediated complex traits Across S. cerevisiae. *PLoS Genet*, 11(1):e1004913, Jan 2015.

[63] D. K. Gonda, A. Bachmair, I. Wünning, J. W. Tobias, W. S. Lane, and A. Varshavsky. Universality and structure of the N-end rule. *J Biol Chem*, 264(28):16700–16712, Oct 1989.

[64] E. S. Johnson, P. C. Ma, I. M. Ota, and A. Varshavsky. A proteolytic pathway that recognizes ubiquitin as a degradation signal. *J Biol Chem*, 270(29):17442–17456, Jul 1995.

[65] R. W. Michelmore, I. Paran, and R. V. Kesseli. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci U S A*, 88(21):9828–9832, Nov 1991.

[66] A. Baryshnikova, M. Costanzo, S. Dixon, F. J. Vizeacoumar, C. L. Myers, B. Andrews, and C. Boone. Synthetic genetic array (SGA) analysis in Saccharomyces cerevisiae and Schizosaccharomyces pombe. *Methods Enzymol*, 470:145–179, 2010.

[67] E. Kuzmin, M. Costanzo, B. Andrews, and C. Boone. Synthetic Genetic Array Analysis. *Cold Spring Harb Protoc*, 2016(4):pdb.prot088807, Apr 2016.

[68] T. Gilon, O. Chomsky, and R. G. Kulka. Degradation signals recognized by the Ubc6p-Ubc7p ubiquitin-conjugating enzyme pair. *Mol Cell Biol*, 20(19):7214–7219, Oct 2000.

[69] T. Sommer and S. Jentsch. A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. *Nature*, 365(6442):176–179, Sep 1993.

[70] F. W. Albert, D. Muzzey, J. S. Weissman, and L. Kruglyak. Genetic influences on translation in yeast. *PLoS Genet.*, 10(10):e1004692, Oct 2014.

[71] F. W. Albert, S. Treusch, A. H. Shockley, J. S. Bloom, and L. Kruglyak. Genetics of single-cell protein abundance variation in large yeast populations. *Nature*, 506(7489):494–497, Feb 2014.

[72] G. Yvert, R. B. Brem, J. Whittle, J. M. Akey, E. Foss, E. N. Smith, R. Mackelprang, and L. Kruglyak. Trans-acting regulatory variation in Saccharomyces cerevisiae and the role of transcription factors. *Nat Genet*, 35(1):57–64, Sep 2003.

[73] C. Brion, S. M. Lutz, and F. W. Albert. Simultaneous quantification of mRNA and protein in single cells reveals post-transcriptional effects of genetic variation. *Elife*, 9, Nov 2020.

[74] M. Gaisne, A. M. Bécam, J. Verdière, and C. J. Herbert. A 'natural' mutation in Saccharomyces cerevisiae strains derived from S288c affects the complex regulatory gene HAP1 (CYP1). *Curr Genet*, 36(4):195–200, Oct 1999.

[75] R. B. Wickner. MKT1, a nonessential Saccharomyces cerevisiae gene with a temperature-dependent effect on replication of M2 double-stranded RNA. *J Bacteriol*, 169(11):4941–4945, Nov 1987.

[76] T. Icho, H. S. Lee, S. S. Sommer, and R. B. Wickner. Molecular characterization of chromosomal genes affecting double-stranded RNA replication in Saccharomyces cerevisiae. *Basic Life Sci*, 40:165–171, 1986.

[77] S. Lutz, C. Brion, M. Kliebhan, and F. W. Albert. DNA variants affecting the expression of numerous genes in trans have diverse mechanisms of action and evolutionary histories. *PLoS Genet*, 15(11):e1008375, 11 2019.

[78] I. M. Ehrenreich, J. Bloom, N. Torabi, X. Wang, Y. Jia, and L. Kruglyak. Genetic architecture of highly complex chemical resistance traits across four yeast strains. *PLoS Genet*, 8(3):e1002570, 2012.

[79] M. Sukalo, A. Fiedler, C. Guzmán, S. Spranger, M. C. Addor, J. N. McHeik, M. Oltra Benavent, J. M. Cobben, L. A. Gillis, A. G. Shealy, C. Deshpande, B. Bozorgmehr, D. B. Everman, E. L. Stattin, J. Liebelt, K. M. Keller, D. R. Bertola, C. D. van Karnebeek, C. Bergmann, Z. Liu, G. Düker, N. Rezaei, F. S. Alkuraya, G. Oğur, A. Alrajoudi, C. A. Venegas-Vega, N. E. Verbeek, E. J. Richmond, O. Kirbiyik, P. Ranganath, A. Singh, K. Godbole, F. A. Ali, C. Alves, J. Mayerle, M. M. Lerch, H. Witt, and M. Zenker. Mutations in the human UBR1 gene and the associated phenotypic spectrum. *Hum Mutat*, 35(5):521–531, May 2014.

[80] X. Li and Y. H. Chang. Amino-terminal protein processing in Saccharomyces cerevisiae is an essential function that requires two distinct methionine aminopeptidases. *Proc Natl Acad Sci U S A*, 92(26):12357–12361, Dec 1995.

[81] R. T. Baker and A. Varshavsky. Yeast N-terminal amidase. A new enzyme and component of the N-end rule pathway. *J Biol Chem*, 270(20):12065–12074, May 1995.

[82] K. Renganaath, R. Cheung, L. Day, S. Kosuri, L. Kruglyak, and F. W. Albert. Systematic identification of cis-regulatory variants that cause gene expression differences in a yeast cross. *Elife*, 9, Nov 2020.

[83] M. Zenker, J. Mayerle, A. Reis, and M. M. Lerch. Genetic basis and pancreatic biology of Johanson-Blizzard syndrome. *Endocrinol Metab Clin North Am*, 35(2):243–253, Jun 2006.

[84] C. S. Hwang, M. Sukalo, O. Batygin, M. C. Addor, H. Brunner, A. P. Aytes, J. Mayerle, H. K. Song, A. Varshavsky, and M. Zenker. Ubiquitin ligases of the N-end rule pathway: assessment of mutations in UBR1 that cause the Johanson-Blizzard syndrome. *PLoS One*, 6(9):e24925, 2011.

[85] J. D. Pedelacq, S. Cabantous, T. Tran, T. C. Terwilliger, and G. S. Waldo. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.*, 24(1):79–88, Jan 2006.

[86] N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer, and R. Y. Tsien. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.*, 22(12):1567–1572, Dec 2004.

[87] S. Kredel, F. Oswald, K. Nienhaus, K. Deuschle, C. R?cker, M. Wolff, R. Heilker, G. U. Nienhaus, and J. Wiedenmann. mRuby, a bright monomeric red fluorescent protein for labeling of subcellular structures. *PLoS ONE*, 4(2):e4391, 2009.

[88] A. Grote, K. Hiller, M. Scheer, R. M?nch, B. N?rtemann, D. C. Hempel, and D. Jahn. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res.*, 33(Web Server issue):W526–531, Jul 2005.

[89] A. Wach, A. Brachat, R. P?hlmann, and P. Philippsen. New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. *Yeast*, 10(13):1793–1808, Dec 1994.

[90] A. L. Goldstein and J. H. McCusker. Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast*, 15(14):1541–1553, Oct 1999.

[91] J. H. Zwolshen and J. K. Bhattacharjee. Genetic and biochemical properties of thialysine-resistant mutants of Saccharomyces cerevisiae. *J Gen Microbiol*, 122(2):281–287, Feb 1981.

[92] R. D. Gietz and R. H. Schiestl. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc*, 2(1):31–34, 2007.

[93] A. C. Ward. Rapid analysis of yeast transformants using colony-PCR. *Biotechniques*, 13(3):350, Sep 1992.

[94] F. Hahne, N. LeMeur, R. R. Brinkman, B. Ellis, P. Haaland, D. Sarkar, J. Spidlen, E. Strain, and R. Gentleman. flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinformatics*, 10:106, Apr 2009.

[95] H. Li and R. Durbin. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14):1754–1760, Jul 2009.

[96] J. S. Bloom, I. M. Ehrenreich, W. T. Loo, T. L. Lite, and L. Kruglyak. Finding the sources of missing heritability in a yeast cross. *Nature*, 494(7436):234–237, Feb 2013.

[97] M. D. Edwards and D. K. Gifford. High-resolution genetic mapping with pooled sequencing. *BMC Bioinformatics*, 13 Suppl 6:S8, Apr 2012.

[98] R. S. Sikorski and P. Hieter. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics*, 122(1):19–27, May 1989.

[99] C. S. Hoffman and F. Winston. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli. *Gene*, 57(2-3):267–272, 1987.

[100] R. M. Horton, H. D. Hunt, S. N. Ho, J. K. Pullen, and L. R. Pease. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, 77(1):61–68, Apr 1989.