Variability in a Short Tandem Repeat Mediates Complex Epistatic Interactions in *Arabidopsis thaliana*

Maximilian Oliver Press and Christine Queitsch¹

University of Washington, Department of Genome Sciences, Seattle, Washington 98195

ORCID ID: 0000-0001-9233-2157 (M.O.P.)

ABSTRACT Short tandem repeats (STRs) are hypervariable genetic elements that occur frequently in coding regions. Their high mutation rate readily generates genetic variation, contributing to adaptive evolution and human diseases. We previously reported that natural *ELF3* polyglutamine variants cause reciprocal genetic incompatibilities in two divergent *Arabidopsis thaliana* backgrounds. Here, we dissect the genetic architecture of this incompatibility, revealing as many as four loci putatively interacting with *ELF3*. We were able to specifically identify one such *ELF3*-interacting gene, *LSH9*. We further used a yeast two-hybrid strategy to identify proteins whose physical interactions with *ELF3* were affected by polyglutamine tract length. We found two proteins for which this was the case, *ELF4* and *AtGLDP1*. Using these two approaches, we identify specific genetic interactions and physical mechanisms by which the *ELF3* polyglutamine tract may mediate the observed genetic incompatibilities. Our work elucidates how STR variation, which is generally underascertained in population-scale sequencing, can contribute to phenotypic variation. Furthermore, our results support our proposal that highly variable STR loci can contribute to the epistatic component of heritability.

KEYWORDS polyglutamine; short tandem repeat; microsatellite; epistasis; robustness

EVOOLUTION is a tinkerer rather than a designer (Jacob 1977; Alon 2003); that is, adaptations are generally short-term, incremental fixes rather than alterations in fundamental biological plans. This principle is believed to underlie many design properties of biological systems. Thus, many (or most) genetic adaptations may be compensations for other genetic variants in a given background (Szamecz *et al.* 2014). One abundant source of genetic variation for such tinkering lies in short tandem repeats (STRs), genetic elements with high mutation rates. Due to these high mutation rates, STRs may be more likely than substitutions to contribute adaptive variants on a per-locus basis (Kashi *et al.* 1997; Gemayel *et al.* 2010; Hannan 2010). If compensation plays an important role in adaptation as predicted, we may expect that STRs also contribute compensatory variants at a higher rate (on average) than other loci. Although it may

be difficult to establish directionality of compensation in any given case, for instance by ordering mutations chronologically, overall STRs should act as compensatory mutations more often than less mutable loci. To sum up, if “tinkering” is a dominant mode of adaptation, variable STRs affecting phenotype are then expected to show frequent epistasis with other loci. Indeed, this expectation is borne out in the handful of well-characterized STRs (Press *et al.* 2014).

One such STR resides in the *Arabidopsis thaliana* gene *ELF3*, where it encodes a polyglutamine tract that varies in length across different natural strains (Tajima *et al.* 2007; Undurraga *et al.* 2012). We have previously shown that these *ELF3*-STR variants have strong effects on phenotype, and that these effects differ depending on the genetic background expressing a particular variant (Undurraga *et al.* 2012). These observations suggest that background-specific variants are modifying the effect of STR alleles through epistasis. The high variability of the *ELF3*-STR relative to expectations suggests that this STR may compensate for many background-specific polymorphisms across globally distributed strains of *A. thaliana*.

ELF3 has been previously identified as a plausible candidate gene underlying a QTL for trait variance (*i.e.*, noise) in the phenotypes of *A. thaliana* recombinant inbred lines

Copyright © 2017 by the Genetics Society of America
doi: 10.1534/genetics.116.193359

Manuscript received July 1, 2016; accepted for publication October 27, 2016;
published Early Online November 17, 2016.

Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.193359/-/DC1.

¹Corresponding author: Foegen S-410, Box 355065, Foegen Bldg., Room S410B,
3720 15th Ave., NE, University of Washington, Seattle, WA 98195-5065. E-mail:
queitsch@uw.edu

(Jimenez-Gomez *et al.* 2011; Lachowiec *et al.* 2015). These results invite comparison to known “robustness genes” such as HSP90 (Sangster *et al.* 2007, 2008a,b), which can reveal or conceal the phenotypic consequences of many other genetic variants. A mechanistic explanation of this robustness phenomenon is epistasis, in which a robustness gene interacts with many other loci (Queitsch *et al.* 2012; Lachowiec *et al.* 2015), as for the promiscuous chaperone HSP90 (Taipale *et al.* 2010). Our previous findings and the many studies describing ELF3’s crucial functions in plant development lead us to hypothesize that *ELF3* lies at the center of an epistatic network and that the ELF3 polyglutamine tract modifies these interactions.

It is well-established that ELF3 functions promiscuously as an adaptor protein in multiple protein complexes that are involved in a variety of developmental pathways (Liu *et al.* 2001; Yu *et al.* 2008; Yoshida *et al.* 2009; Nusinow *et al.* 2011; Chow *et al.* 2012; Huang *et al.* 2016). Polyglutamine tracts such as the one encoded by the *ELF3*–STR often mediate protein interactions (Perutz *et al.* 1994; Stott *et al.* 1995; Schaefer *et al.* 2012). Therefore, it is plausible that variation in the ELF3 polyglutamine tract affects ELF3’s interactions with its partner proteins. The ELF3 C terminus, which contains the STR-encoded polyglutamine tract, is necessary for nuclear localization (Herrero *et al.* 2012) and ELF3 homodimerization (Liu *et al.* 2001), but thus far relatively few proteins (PIF4 and TOC1) have been shown to interact with this domain (Nieto *et al.* 2014; Huang *et al.* 2016). Thus, the phenotypic and epistatic effects of *ELF3*–polyQ variation may arise from altered protein interactions, altered ELF3 nuclear localization, altered regulation of the PIF4 developmental integrator, or a combination thereof. In consequence, there are a variety of potential mechanisms for epistatic interactions of *ELF3* and other loci, and thus we consider the broadest interpretation of epistasis: the simple nonadditivity of effects of genetic variants.

Here, we use genetic analysis of the *ELF3*–STR-based incompatibility to identify *ELF3*–STR-interacting loci, and describe how *ELF3*–polyQ variation affects ELF3 protein interactions. This work presents evidence that this STR anchors a complex network of epistasis, potentially due to its role as a compensatory modifier of several other loci.

Materials and Methods

Plant material and growth conditions

We generated a Columbia (Col) × Wassilewskija (Ws) F₂ population by crossing to genetically map loci interacting with *ELF3* in contributing to hypocotyl length. Seedling hypocotyl length was assayed in seedlings grown for 15 days in incubators set to short days (8 hr light:16 hr dark) at 22° on vertical plates as described previously (Undurraga *et al.* 2012). The *elf3-200* (Undurraga *et al.* 2012) and *elf3-4* (Hicks *et al.* 1996) mutants have been previously described. Transfer DNA (T-DNA) lines (Alonso *et al.* 2003; Kleinboelting *et al.* 2012) were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus,

OH). After phenotypic analysis, we stored seedlings at 4° before genotyping.

Genotyping

We genotyped the *ELF3*–STR across 720 F₂ segregants to identify *ELF3* homozygotes for further genomic analysis. The STR differs by 27 bp and is thus readily distinguishable by agarose gel electrophoresis. We used one to two true leaves from each seedling for DNA extraction. PCR was performed in 10-μl volume containing 0.5 μM primers (Supplemental Material, Table S1), 0.2 μM each deoxynucleotide triphosphate (dNTP), 1 μl 10× ExTaq buffer, and 0.1 units ExTaq (Takara, Tokyo, Japan); with initial denaturation step of 95° for 5 minutes, followed by 40 cycles of 95° 30 seconds, 49° 20 seconds, 72° 10 seconds, with a final extension step at 72° for 5 minutes. For PCR analysis of other loci, PCR was performed in 20 μl volume containing 0.5 μM primers (Table S1), 0.2 μM each dNTP, 2 μl 10× ExTaq buffer, and 0.25 units Taq polymerase (NEB, Ipswich, MA); with initial denaturation step of 95° for 5 minutes, followed by 35 cycles of 95° 30 seconds, 55° 30 seconds, 72° 1 minute, with a final extension step at 72° for 5 minutes.

Genome resequencing

We selected *ELF3* homozygotes among the F₂’s for genotyping by sequencing according to their phenotypes. For each *ELF3*–STR allele, we selected 24 *ELF3* homozygotes (Ws/Ws and Col/Col) at each phenotypic extreme (the shortest and longest hypocotyls). The sampling of extremes is an effective and statistically justified method for genetic mapping (Lander and Botstein 1989). These 96 individuals were analyzed in a genotyping-by-sequencing approach (Figure S1 and Table S3). These seedlings were transplanted to soil and grown under long days for 2–3 weeks. They were then stored at 4° until DNA extraction was performed. One late rosette-stage Ws individual was used for Ws whole-genome resequencing. DNA extraction was performed using the DNeasy Plant Mini kit (QIAGEN, Valencia, CA) according to the kit protocol. This DNA was quantified using high-sensitivity Qubit fluorescence analysis (ThermoFisher Scientific, Waltham, MA) and resequenced with *ELF3*–STR primers (Table S1). We used 10 ng DNA from each F₂ segregant in NextEra transposase library preparations (Illumina, San Diego, CA), or a standard 50-ng preparation for the Ws library. Library quality was assessed on a BioAnalyzer (Agilent, Santa Clara, CA) or agarose gels. The Ws individual was sequenced in one 300-cycle MiSeq v2 run (300 bp single-end reads) to ~12× coverage. The F₂ segregant libraries were pooled and sequenced in one 200-cycle HiSeq v3 run to ~2× average coverage (100-bp paired-end reads, Table S3).

Sequence analysis

We aligned reads to the Col reference genome using BWA v0.7.5 MEM (Li 2013), and variants were called using SAMtools v0.1.19 (Li *et al.* 2009). We identified high-quality Ws variants (*Q* > 40) from Ws parent data and compared these with variants in previously sequenced related strains (Gan *et al.* 2011).

We combined all F_2 segregant genotype calls into a single variant call format (VCF) file and filtered these for loci with such Ws variants. We used SNPtools (Wang *et al.* 2013) to perform haplotype and genotype imputation for each locus in F_2 segregants. For workflows employed in sequence analysis, see File S1. Following sequence analysis, we found one individual to be a heterozygote at the *ELF3* locus. We omitted this individual from all following analyses requiring *ELF3* homozygotes.

QTL analysis

We reduced F_2 genotypes to a set of 500 loci randomly sampled from the imputed genotypes, plus a single nucleotide variant (SNV) marking the *ELF3* locus. We used these genotypes to estimate a genetic map and perform QTL analysis using the R/qtl package (Broman *et al.* 2003). We genotyped 30 additional F_2 's (further *ELF3* homozygotes with extreme phenotypes) with PCR markers (Table S1) to increase QTL accuracy. To directly test for epistasis with *ELF3*, we adapted a previously described method for evaluating interactions with prespecified loci (Sangster *et al.* 2008a). We defined an empirical null distribution for the difference of LOD scores ("delta LOD") expected between random samples of genotyped F_2 's, comparing this to the actual difference observed between *ELF3-Col* and *ELF3-Ws* homozygotes. We implemented this custom nonparametric epistasis test using R/qtl functions. For a more detailed description of commands and the epistasis test, see File S1.

Candidate gene analysis

We assessed candidate gene function and interaction with *ELF3* using homozygous T-DNA insertion lines (Alonso *et al.* 2003; Kleinboelting *et al.* 2012) obtained from the *Arabidopsis* Biological Resource Center. We phenotyped these lines for hypocotyl length under SD at 15 days as for F_2 's. All such experiments were performed at least twice. We obtained *elf3 lsh9* and *elf3 nup98* double mutants by crossing relevant lines and genotyping (primers in Table S1). Expression analysis confirmed no detectable *LSH9* expression in the *lsh9* mutant, suggesting that it is a null mutant. Mutant lines are listed in Table S2.

Yeast two hybrid

We used yeast two hybrid (Y2H) to assess different proteins for interactions with *ELF3* and specifically to test the hypothesis that *ELF3*-polyQ variation affects these interactions. We PCR cloned *ELF3* variants with different STR lengths out of complementary DNAs (cDNAs) of previously described *A. thaliana* carrying *ELF3* transgenes (Undurraga *et al.* 2012) into the *XmaI/BamHI* sites of pGBKT7. These included synthetic *ELF3*s with 0Q (no polyQ), 7Q (endogenous variant in Col), 16Q (endogenous variant in Ws), and 23Q (endogenous to strains Br-0 and Bur-0). We PCR cloned genes to be tested for *ELF3* interactions into the *EcoRI/XhoI* sites of pGADT7 from cDNAs of indicated strains (Table S1 for primers). Clones were confirmed by restriction digest and

sequencing. We performed a Y2H screen of the *ELF3*-7Q variant against the *Arabidopsis* Mate and Plate cDNA library (Clontech, Madison, WI), essentially according to the manufacturer's instructions, with selections performed on Synthetic Complete (C) –Leu –Trp –His plates incubated at 23°. We tested clones, which also showed activation of the *ADE2* reporter gene (and did not autoactivate) against the various *ELF3*-polyQ constructs (see File S1 for details, full details on clones given in File S2).

As a more quantitative assay of Y2H interaction, we assayed LacZ activity through X-gal cleavage essentially as previously described (Möckli and Auerbach 2004), again in strains using PJ69-4 α as *MAT α* parent. LacZ expression is, like *HIS3* and *ADE2*, also driven by Y2H in this system. For weakly activating constructs (GLDP1 and *ELF4*), we used 0.2 absorbance units of yeast in each assay to reduce background and assessed color development at points between 16 and 72 hr of incubation at room temperature.

Quantitative PCR

For measuring *ELF3* and *LSH9* transcript levels, we collected ~30 mg pooled aerial tissue from short-day-grown seedlings of each relevant genotype at Zeitgeber time (ZT) 8 at 7 days postgermination. RNA preparation, reverse transcription, and quantitative PCR (qPCR) was performed as described previously (Undurraga *et al.* 2012), using primers in Table S1.

Statistical analysis

All statistical analyses and plotting was performed using R 3.2.1 (R Development Core Team 2016). Analysis scripts are provided with data as detailed below.

Data availability

High-throughput sequencing data are available in BAM format at National Center for Biotechnology Information Sequence Read Archive accession no. SRP077615. Processed genotype data, phenotype data, and analysis code are available at https://figshare.com/articles/Variable_ELF3_polyQ_modules_complex_epistasis_data_and_code/3467717.

Results

Genetic analysis of *ELF3*-STR effects on hypocotyl length

To investigate the genetic architecture of epistasis for the *ELF3*-STR, we used hypocotyl length under short days, a trait controlled by *ELF3*-STR variation (Undurraga *et al.* 2012). We previously reported a mutual incompatibility of these *ELF3*-STR variants between the Col and Ws strains using transgenic analysis. Here, we phenotyped a Col $\times$ Ws F_2 population for this trait, which we expected to show substantial variation according to the segregation of the *ELF3*-STR variants.

The Col and Ws strains did not substantially differ in this trait ($P = 0.16$, Kolmogorov-Smirnov test, Figure 1A). Although most F_2 seedlings showed phenotypes within the

range of the two parental lines, the F_2 phenotypic distribution showed a long upper tail of transgressive variation and consequently a different distribution from either parent ($P = 0.0039$ against Ws, $P = 0.055$ against Col, Kolmogorov–Smirnov tests). As longer hypocotyls in light conditions indicate *ELF3* dysfunction in the circadian clock (Liu *et al.* 2001), this observation agrees with the hypothesized incompatibility of Col and Ws alleles (which include the *ELF3*–STR). We replicated this observation in a much larger population (1106 seedlings), which was used for further genetic analysis.

To investigate the genetic basis of the phenotypic transgression in hypocotyl length, we harvested approximately two of three seedlings representing the longest and shortest hypocotyls (Figure S1). We genotyped the *ELF3*–STR in these seedlings and observed a strong main effect of *ELF3* on phenotype (Figure 1B), in which the Col allele of *ELF3* frequently showed transgressive phenotypes, although some individuals homozygous for the Ws allele also showed transgressive phenotypes. We cannot exclude any effect from *ELF3*–linked variants other than the STR, but given the known effects of *ELF3*–STR variation in these backgrounds (Undurraga *et al.* 2012), we consider the STR as the most plausible Chr2 variant underlying the observed transgression. Specifically, a naïve regression analysis of the data in Figure 1B indicated that each *ELF3*–Col allele increased hypocotyl length by 0.87 ± 0.077 mm, and that the *ELF3* locus thereby explained 15% of phenotypic variation. This analysis is of course misleading, because it implies that Col seedlings should show longer hypocotyls than Ws seedlings due to *ELF3* genotype, which is not the case (Figure 1A). Moreover, this observation suggests that the relevant epistatic model under which *ELF3* operates with other loci is reminiscent of capacitance (Bergman and Siegal 2003), in which mutational effects are revealed by *ELF3* dysfunction.

Among these seedlings, the individuals with extreme phenotypes and individuals homozygous at the *ELF3* locus are expected to be most informative about *ELF3*–STR effects on phenotype. Furthermore, *ELF3* genetic interactions are expected to be most apparent in *ELF3* homozygotes. Therefore, we focused our analysis on these homozygotes, which we analyzed using a genotyping-by-sequencing approach (Figure S1). For details of our approach, see *Materials and Methods* and Figure S2, Figure S3).

With these data, we performed a one-locus QTL scan to identify chromosomal regions contributing to hypocotyl length (Figure 2A). This analysis indicated a QTL on Chr2 corresponding to *ELF3* as expected, but also significant QTL on Chr1, Chr4, Chr5, and possibly one or more additional QTL on Chr2 affecting the phenotype. A two-dimensional QTL scan suggested that at least some of these QTL interact epistatically with the *ELF3* locus (Figure S4).

We investigated potential epistatic effects of *ELF3* on other loci, performing one-dimensional QTL scans on each *ELF3* homozygote background separately (masking the genotypes of all individuals with other *ELF3* genotypes). This analysis identified loci whose effects are potentially contingent on one *ELF3* allele or the other. We observed that the same LOD

Figure 1 Phenotypic transgression in Col $\times$ Ws F_2 segregants. (A) F_2 segregant phenotypes compared to parents and an *elf3* null mutant in the Col background; $n \approx 50$ for each homozygous line and $n \approx 100$ for F_2 's. Colors indicate genotypes as indicated. Hypocotyl length was determined at 15 days under short days. (B) Phenotypic distributions of a large population of Col $\times$ Ws F_2 segregants, for 720 extreme individuals genotyped at the *ELF3* locus. $N = 1106$ total seedlings. A total of 386 seedlings with intermediate phenotypes were not genotyped (indicated by the black box).

peaks were replicated well in *ELF3*–Col homozygotes, but poorly in *ELF3*–Ws homozygotes (Figure 2B). Notably, a second Chr2 QTL was thus made obvious in this analysis, indicating that loci other than *ELF3* on Chr2 are relevant to the phenotype (at least in *ELF3*–Col plants). These results suggested that the *ELF3*–STR genotype is epistatic to at least three other loci controlling hypocotyl elongation, whose effects are masked in *ELF3*–Ws plants. To formally test for epistasis with *ELF3*, we compared the difference in LOD scores between the two QTL scans at all loci to a permutation-based null distribution, and found that the peaks on Chr1 and Chr4 (and to a lesser extent Chr2) were all stronger in the *ELF3*–Col background (Figure 2C). Consequently, each of these QTL likely contains one or more background-specific *ELF3* interactors.

We considered the genetic contribution of these loci to the phenotype using a multiple QTL mapping approach, using both the independently estimated QTL locations and a refined model reestimating QTL positions based on information from all QTL (Table S4). In each case, loci on Chr1, Chr2, Chr4, and Chr5 were supported, along with interactions between Chr2 (*ELF3*) and the other two loci on Chr1 and Chr4. In the refined model, the *ELF3* locus showed a dramatically higher LOD score and other loci generally decreased in importance (especially the second Chr2 locus), again suggesting epistatic masking by *ELF3*. We conclude that although *ELF3* interacts epistatically with a variety of other loci in determining hypocotyl length, the principal contributors to *ELF3*–mediated effects on the trait appear to be on Chr1 and Chr4. Moreover, direct inspection of phenotypic effects of *ELF3* in interaction with each putative locus among F_2 segregants supported the hypothesis of epistasis with *ELF3* most clearly for the Chr1 and Chr4 loci (Figure S5).

Candidate gene analysis identifies *LSH9* as a genetic interactor of *ELF3*

The chromosome intervals identified by our QTL analysis encompassed a large number of genes and partially overlapped with an inversion between these backgrounds on Chr4 (Rowan

Figure 2 QTL analysis identifies interactions of *ELF3* with multiple loci. (A) One-dimensional QTL scan including all sequenced F_2 's. Horizontal line indicates 99% significance threshold based on permutations. (B) QTL scan stratified by *ELF3*-STR genotype (all genotypes but those of indicated F_2 's masked in each analysis). Horizontal lines of each color indicate 99% significance threshold based on permutations for each scan. (C) A nonparametric test of epistasis between *ELF3* and other loci, using the independent QTL scans shown in B, *ELF3* is located on chromosome 2. Horizontal lines indicate 95% significance threshold based on permutations.

taxis between *ELF3* and other loci, using the independent QTL scans shown in B, *ELF3* is located on chromosome 2. Horizontal lines indicate 95% significance threshold based on permutations.

et al. 2015), although the QTL peak was >1 Mb outside the mapped inversion. Previous work using a multiparent *A. thaliana* mapping population (Kover *et al.* 2009) also identified possible candidate genes affecting hypocotyl length in these regions (Khattak 2014).

We phenotyped mutants of several candidate genes in the Col background under the conditions of our intercross experiment (15-day SD hypocotyl length, Figure S6). We observed small phenotypic effects of the T-DNA insertion mutants *lsh9* and *nup98*. However, these small effects on their own cannot explain the transgressive phenotypic variation in F_2 's (Figure 1A).

The secondary Chr2 peak (other than *ELF3*) included the known *ELF3* interactor *PIF4*. *pif4* null mutants are known to partially suppress the long hypocotyl defect of *elf3* mutants, and thus *PIF4* is a candidate gene with prior support. However, our genetic data showed comparatively weak signal at the Chr2 peak when accounting for the *ELF3* QTL, and its interval was very wide (18Mb, Table S4). When examining the actual phenotypic effects of this QTL, the interaction appeared ambiguous compared to the Chr1 and Chr4 signals (Figure S5). We therefore focused on the novel candidate genes *NUP98* on Chr1 and *LSH9* on Chr4, given the strong support for their interaction with *ELF3* in this population (Figure 2, Figure S5, and Table S4).

We generated double mutants between mutants of these genes and the *elf3* null mutant to determine whether these genes interacted epistatically with *ELF3*. We found little evidence for an interaction between *nup98* and *elf3* mutations (Figure S6C). However, we detected a significant interaction between *ELF3* and *LSH9*, in the form of reciprocal sign epistasis between the two null mutants affecting hypocotyl length (Figure 3A). Although *lsh9* single mutants had significantly shorter hypocotyls than WT, *lsh9 elf3* double mutant hypocotyls were substantially longer than in *elf3* single mutants. *LIGHT-DEPENDENT SHORT HYPOCOTYLS 9* (*LSH9*) is an uncharacterized gene belonging to a gene family named for *LSH1*, which is known to act in hypocotyl elongation (Zhao *et al.* 2004). Publicly available expression data indicated that *LSH9* transcript levels are highest in the hypocotyl and root and show light-dependent circadian fluctuations (Winter *et al.* 2007). Like other genes in this family, *LSH9* encodes a putative nuclear localization sequence but has no other distinguishing features.

To test our hypothesis that *ELF3*-STR-mediated epistasis may be due to altered protein interactions, we investigated whether *LSH9* and *ELF3* interacted physically using Y2H. However, we were unable to detect a physical interaction between the Col or Ws variants of *LSH9* and *ELF3* (Figure S7), suggesting a different mechanistic basis for the observed genetic interaction. For example, *LSH9* expression may depend on *ELF3* function as a transcriptional regulator. Alternatively, *ELF3* expression may depend on *LSH9* function. We tested both hypotheses by measuring expression levels of *LSH9* and *ELF3* in, respectively, *elf3* or *lsh9* mutant backgrounds (*elf3* mutants were available in both Col and Ws backgrounds, but *lsh9* mutants were available only in Col). *ELF3* expression levels were unchanged in *lsh9* mutants (Figure S8). Moreover, levels of *LSH9* transcript did not significantly differ between WT and *elf3* mutants in either strain background. However, *LSH9* expression was reduced in the Ws background relative to Col independently of *ELF3* genotype (Figure 3B). This result is consistent with the observed phenotypic interaction in F_2 's, which showed elongated hypocotyls when Col alleles at the *ELF3* locus cosegregated with Ws alleles at the *LSH9* locus (Figure S5), thereby pairing poorly functioning *ELF3* alleles with potentially lower *LSH9* expression levels.

Taken together, *ELF3*-*LSH9* epistasis between Col and Ws may be due to regulatory changes between these two backgrounds altering *LSH9* transcript levels. For example, genomic data indicate that the *LSH9* promoter contains an STR polymorphism in the Ws background that may alter *LSH9* expression (Alonso-Blanco *et al.* 2016).

***ELF3*-polyQ tract variation affects known protein interactions**

While our QTL analysis identified several potential interactors of the *ELF3* locus, there was no obvious mechanistic interpretation of *ELF3* polyQ variation's effect on protein function. Based on prior knowledge on *ELF3* and polyQ tract function (Schaefer *et al.* 2012; Huang *et al.* 2016), we hypothesized that polyQ variation may affect *ELF3*'s interactions with other proteins. Thus, in parallel with our genetic analysis, we used Y2H to directly identify *A. thaliana* proteins whose physical interactions with *ELF3* are polyQ modulated. We first explored whether several variants of *ELF3* (including 7Q and 16Q) show Y2H interactions with the well-described *ELF3* interactors PHYB

Figure 3 *ELF3* interacts genetically with *LSH9*, which shows background-specific expression. (A) Double mutant analysis of *elf3* and *lsh9* seedlings (in the Col background) grown for 15 days under short days. ANOVA analysis of the interaction between the mutant effects on phenotype is displayed. Error bars indicate SD, $n > 35$ for each genotype. Experiments were repeated with similar results. (B) qRT-PCR analysis of *LSH9* transcript levels across genotypes. Seedlings were grown under SD and harvested at ZT8 (dusk). *LSH9* expression is expressed as a proportion of Ws expression, normalized relative to *UBC21*, error bars are SE from four biological replicates.

(Liu *et al.* 2001), *ELF4* (Nusinow *et al.* 2011; Herrero *et al.* 2012) (Figure 4), and *PIF4* (Nieto *et al.* 2014). None of the *ELF3* constructs showed autoactivation in yeast when paired with an empty vector (Figure S9). The *ELF3*-interacting domain of *PHYB* has two coding variants between Col and Ws, and we thus tested both Ws and Col variants of this domain. We found that both forms showed apparently equal affinity with all polyQ variants of *ELF3*. *ELF4*, which has no coding variants between Col and Ws, also interacted with all polyglutamine variants of *ELF3*, although rather weakly compared to *PHYB*. Under these conditions, a subtle preference of *ELF4* for some polyQ variants was apparent (e.g., *ELF3*-16Q vs. *ELF3*-0Q). We confirmed this preference in a quantitative Y2H LacZ assay (Figure 4B). We were not able to replicate the previously reported *ELF3*-*PIF4* interaction (Nieto *et al.* 2014) for any *ELF3*-polyQ variant in our Y2H system (Figure S10) and were thus unable to evaluate effects of polyQ variation on *ELF3*-*PIF4* interactions.

Together, our data suggest that *ELF3*-polyQ tract variation can affect *ELF3* protein interactions, in particular if these interactions are weaker (as for *ELF4*) and presumably more sensitive to structural variation in *ELF3*.

Y2H screen identifies three novel *ELF3* interactors, one of which is polyQ modulated

None of the known *ELF3* interactors were encoded by genes located in the major Chr1 and Chr4 QTL identified by our genetic screen. If the *ELF3*-polyQ tract mediates protein

Figure 4 Y2H interaction of *ELF3* with known protein interactors can be modulated by polyQ variation. (A) Yeast carrying indicated vectors were spotted in fivefold dilutions onto C –Leu –Trp (C–LT) or C –Leu –Trp –His –Ade (C–LTHA) media. *PHYB*-Cterm, previously defined C-terminal truncations of *PHYB* sufficient for *ELF3* interaction (Liu *et al.* 2001) from the Col and Ws backgrounds. For each protein X, experiments were repeated with independent PJ69-4α + pGADT7-X transformants with similar results. (B) LacZ assays support polyQ effects on *ELF3*-*ELF4* interaction. The strains shown in A also express *LacZ* from the Y2H promoter, whose activity was assayed in yeast cell lysates (see Materials and Methods). In each assay, all observations are expressed relative to the activity of the empty vector, whose mean is set to 0. Error bars indicate SD across three technical replicates. This experiment was repeated with similar results.

interactions, these regions should contain additional, previously undescribed polyQ-modulated *ELF3* interactors. We screened the *ELF3*-7Q protein for interactions with proteins from a commercially available library derived from Col, to detect *ELF3*-protein interactions within the Col background.

We subjected Y2H positives to several rounds of confirmation (Supplementary Material), yielding a total of three novel proteins that robustly interacted with *ELF3*: PLAC8 domain-containing protein AT4G23470, LUL4, and AtGLDP1 (Figure 5). AT4G23470 was recovered in two independent clones, and LUL4 was recovered in three independent clones. The PLAC8 domain protein AT4G23470 is encoded by a gene within the QTL interval on chromosome 4, but this protein showed no variation in affinity among the various *ELF3*-polyQs. LUL4, a putative ubiquitin ligase, is not encoded in any of the mapped QTL and also shows no variation in affinity among the various *ELF3*-polyQs. Thus, differential interaction with these proteins is unlikely to underlie the observed epistasis.

In contrast, the AtGLDP1 protein, which is encoded on Chr4 but not within the QTL interval, appeared to show a subtle preference for the synthetic *ELF3*-0Q construct over longer polyQ tracts. We confirmed this preference in a quantitative LacZ assay (Figure 5B). Although our screen is unlikely to exhaust hitherto-unknown *ELF3* interactors, our data suggest that the *ELF3*-polyQ tract can affect *ELF3*'s interactions with other proteins. Moreover, polyQ variation appears to affect weaker *ELF3*-protein interactions; strong protein interactions (for example *ELF3*-LUL4 and *ELF3*-*PHYB*) are robust to polyQ variation.

Figure 5 Y2H screen identifies new interactors of ELF3. (A) Yeast carrying indicated vectors were spotted in fivefold dilutions onto C –Leu –Trp (C–LT) or C –Leu –Trp –His –Ade (C–LTHA) media. For each protein X, experiments were repeated with at least two independent PJ69-4α + pGADT7-X transformants with similar results. (B) LacZ assays support polyQ effects on ELF3–AtGLDLP1 interaction. The strains shown in A also express *LacZ* from the Y2H promoter, whose activity was assayed in cell lysates (see *Materials and Methods*). In each assay, all observations are expressed relative to the activity of the empty vector, whose mean is set to 0. Error bars indicate SD across three technical replicates. This experiment was repeated with similar results.

Discussion

The contribution of STR variation to complex traits is thought to be considerable (Kashi *et al.* 1997; Press *et al.* 2014). Specifically, it has been proposed that STR variation contributes disproportionately to the epistatic term of genetic variance, due to its potential to contribute compensatory mutations. However, the molecular mechanisms by which different STRs contribute to genetic variance should derive from their particular features. For instance, polyQ variation may be expected to affect protein interactions (Perutz *et al.* 1994; Schaefer *et al.* 2012) or the transactivation activity of affected proteins (Escher *et al.* 2000). In this study, we considered the case of the previously described *ELF3*–STR (Undurraga *et al.* 2012).

We found that the genetic architecture of *ELF3*–dependent phenotypes is highly epistatic between the divergent Col and Ws strains, leading to substantial phenotypic transgression in the well-studied hypocotyl length trait (Zagotta *et al.* 1996; Dowson-Day and Millar 1999; Nusinow *et al.* 2011). We identified at least three QTL showing genetic interactions with the *ELF3*–STR in a Col × Ws cross. These QTL generally did not coincide with obvious candidate genes known to affect *ELF3* function (with the potential exception of *PIF4*). Our confirmation of one genetic interaction (*LSH9*) in the Col background suggests that these QTL encompass variants affecting hypocotyl length in tandem with *ELF3*–STR variation. We cannot formally exclude the hypothesis that variants linked to *ELF3* (other than the Col and Ws *ELF3*–STR variants) may contribute to the observed phenotypic variation. For example, the *ELF3*–A362V substitution in the *A. thaliana* strain Sha

affects *ELF3* function in the circadian clock (this site is invariant between Col and Ws) (Anwer *et al.* 2014). However, our previous work demonstrated that *ELF3*–STR variation suffices to produce strong phenotypic incompatibility between the Col and Ws background (Undurraga *et al.* 2012). Therefore, we reason that *ELF3*–STR variation is the most parsimonious explanation for the phenotypic variation, and in particular the observed transgression in the Col × Ws cross.

We further used Y2H screening to explore whether *ELF3* polyQ tract variation affects protein interactions. *ELF3*’s promiscuous physical associations with other proteins are essential to its many functions in plant development (Liu *et al.* 2001; Kolmos *et al.* 2011; Nusinow *et al.* 2011; Herrero *et al.* 2012). Disruption of these interactions suggested an attractive mechanism by which *ELF3*–polyQ tract variation might affect *ELF3* function. Assaying several known and novel *ELF3*–interacting proteins yielded evidence for a modest effect of polyQ variation on weaker protein interactions. However, there was no generic requirement for specific *ELF3* polyQ tract lengths across all interactors. Indeed, the modest effects that we found were interactor specific and thus not likely to generalize.

We did find that the *ELF3*–*ELF4* interaction, which is crucial for circadian function and thus hypocotyl length (Nusinow *et al.* 2011; Herrero *et al.* 2012), demonstrates a subtle preference for the Ws 16Q–*ELF3* variant. However, there is no sequence variation in *ELF4* between Col and Ws and we did not detect the *ELF4* locus by QTL analysis, suggesting that this binding preference does not explain the transgressive phenotypes revealed by *ELF3*–STR variation (Figure 1A). However, the subtle polyQ dependence of the *ELF3*–*ELF4* interaction may play a role through indirect interactions.

Alternatively, rather than modulating *ELF3* function as an encoded polyQ tract, the *ELF3*–STR may affect *ELF3* transcription or processing. A previous study of an intronic STR in *A. thaliana* demonstrated that certain hyperexpanded STR alleles led to dysregulation of the *III1* gene (Sureshkumar *et al.* 2009), presumably due to aberrant processing of *III1* transcripts. Others have previously argued that such “informational” (as opposed to “operational”) processes are more likely to have genetic or physical interactions (Jain *et al.* 1999). We have not tested this hypothesis, although our previous studies found no correlation between *ELF3*–STR variation and *ELF3* expression across many natural strains (Undurraga *et al.* 2012).

Taken together, our findings support a model in which highly variable STRs can contribute to the epistatic component of heritability through both direct and indirect functional interactions with other loci. One potential concern with our model is that one might expect highly epistatic loci to be more constrained than other loci, and thus unlikely to accumulate more compensatory mutations. Nonetheless, the existence of the *ELF3*–STR epistatic network suggests that the constraints are not insurmountable, potentially as a result of a cascade of compensations. The underlying evolutionary dynamics of this process, and of high mutation rates in general, deserve more study.

ELF3 is a model for subexpansion polyQ variation

PolyQ variation is best known from hyperexpansions that are associated with several incurable human neurological disorders (Orr and Zoghbi 2007; Fondon *et al.* 2008; Usdin 2008; Hannan 2010), in which CAG (although generally not CAA) repeats dramatically elongate to lengths >50 units, reaching over 200 units in some patients. We argue that these hyperexpansion disorders are poor models for the functional impact of subexpansion variation of polyQ tracts and propose instead that ELF3–polyQ might serve as a more appropriate model.

Although ELF3–polyQ tract variation can reach a length associated with disease in humans, it differs qualitatively from the well-studied human polyQ tract hyperexpansions. Human polyQ hyperexpansions are associated with protein aggregation or plaque formation, a phenomenon that requires a sufficiently long uninterrupted polyglutamine domain (Sharma *et al.* 1999; Lu and Murphy 2015). These previous *in vitro* studies suggest that although the ELF3–polyQ variants in Col and Ws are sufficiently different to alter phenotype, neither is long enough to lead to aggregation by these mechanisms (although the ELF3–23Q variant may be within this range).

Next, the effects of the disease-associated polyQ hyperexpansions are generally dominant, due to the nature of their molecular effects, which are generally thought to show a protein gain of function (Orr and Zoghbi 2007; Fondon *et al.* 2008). We observed no evidence that ELF3–polyQ variation behaves in a dominant fashion, but rather that Col × Ws F₁'s show approximately WT phenotypes. The effects of ELF3–polyQ variation manifest only when separated from a favorably interacting genetic context (as in segregating F₂'s).

Last, the (sometimes implicit) expectation from polyQ hyperexpansion phenotypes is that there is a linear, or at least monotonic, association between phenotypes and polyQ length. For instance, the degree of huntingtin polyQ expansion is strongly correlated with Huntington's disease severity (Andrew *et al.* 1993). In contrast, there is little evidence for any monotonic relationship between a ELF3–polyQ-related phenotype and ELF3–polyQ length (Undurraga *et al.* 2012; Press *et al.* 2014, 2016). Indeed, all indications are that the mapping between ELF3–polyQ tract length and phenotype is nonmonotonic and strongly contingent on genetic background, unlike the classic polyQ disease models. Our results suggest that subexpansion polyQ tract variation can engage in multiple genetic interactions and at least in some cases modulate protein interactions. More work is needed to evaluate the generality of our findings and determine the breadth of molecular mechanisms by which modest, subexpansion polyQ variation can affect phenotype.

ELF3 as a robustness gene

Here, we operated under the assumption that a few strong polyQ-modulated interactions with ELF3 explain the polyQ-

dependent genetic architectures. Alternatively, the ELF3 polyQ tract may modulate many transient interactions that are perturbed by hypomorphic ELF3 activity. In this interpretation, *ELF3* acts as a “robustness gene” (Lempe *et al.* 2013). The best-described example of such is the protein chaperone HSP90 (Rutherford and Lindquist 1998), whose multiple transient interactions with many proteins—~10% of the yeast proteome (Zhao *et al.* 2005)—lead to pleiotropic effects upon HSP90 inhibition or dysregulation (Sangster *et al.* 2007). ELF3 has been previously proposed as a robustness gene (Jimenez-Gomez *et al.* 2011), consistent with its promiscuity in protein complexes and the pleiotropic nature of *elf3* mutant phenotypes. Our finding that functional modulation of ELF3 by polyQ variation reveals several genetic interactors is consistent with this interpretation.

A similar hypothesis is that ELF3 “gates” robustness effects from robustness genes with which it interacts. For instance, we have recently shown that ELF3 function is epistatic to some of HSP90's pleiotropic phenotypic effects (M. Zisong, P. Rival, M. Press, C. Queitsch, and S. Davis, unpublished data), and ELF4 has also been proposed as a robustness gene governing circadian rhythms and flowering (Lempe *et al.* 2013). Here, we show that polyQ variation affects ELF3–ELF4 binding, which would provide a mechanistic link between ELF3–polyQ effects and a known robustness gene.

These hypotheses remain speculative in the absence of more explicit tests. Nonetheless, we suggest that the pleiotropic effects of polyQ variation in ELF3 (or similar cases) may be better understood by considering *ELF3* as a robustness gene, in which phenotypic effects are determined by a variety of important but individually small interactions of this highly connected epistatic hub.

Acknowledgments

We thank Karla Schultz and Katie Uckele for technical assistance, Choli Lee and Jay Shendure for assistance with high-throughput sequencing of the Col × Ws F₂ population, members of the Shendure laboratory for advice regarding library preparation, Amy Lanctot for generating the pGBK–ELF3–0Q and pGBK–ELF3–23Q constructs, Daniel Melamed and Stanley Fields for guidance in carrying out Y2H experiments and the generous gift of yeast strains, Giang Ong and Maitreya Dunham for access to the MiSeq instrument for resequencing the Ws genome, Stanley Fields and Evan Eichler for access to LightCycler instruments, Elhanan Borenstein and members of the Queitsch and Borenstein laboratories for helpful conversations, and two anonymous reviewers for helpful suggestions toward improving the text. M.O.P. was supported in part by National Human Genome Research Institute Interdisciplinary Training in Genome Sciences grant 2T32HG35-16. C.Q. is supported by National Institutes of Health New Innovator Award DP2OD008371.

Literature Cited

- Alon, U., 2003 Biological networks: the tinkerer as an engineer. *Science* 301: 1866–1867.
- Alonso, J. M., A. N. Stepanova, T. J. Leisse, C. J. Kim, H. Chen *et al.*, 2003 Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657.
- Andrew, S. E., Y. P. Goldberg, B. Kremer, H. Telenius, J. Theilmann *et al.*, 1993 The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* 4: 398–403.
- Anwer, M. U., E. Boikoglou, E. Herrero, M. Hallstein, A. M. Davis *et al.*, 2014 Natural variation reveals that intracellular distribution of ELF3 protein is associated with function in the circadian clock. *eLife* 3: e02206.
- Bergman, A., and M. L. Siegal, 2003 Evolutionary capacitance as a general feature of complex gene networks. *Nature* 424: 549–552.
- Broman, K. W., H. Wu, S. Sen, and G. A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19: 889–890.
- Chow, B. Y., A. Helfer, D. A. Nusinow, and S. A. Kay, 2012 ELF3 recruitment to the PRR9 promoter requires other evening complex members in the *Arabidopsis* circadian clock. *Plant Signal. Behav.* 7: 170–173.
- Dowson-Day, M. J., and A. J. Millar, 1999 Circadian dysfunction causes aberrant hypocotyl elongation patterns in *Arabidopsis*. *Plant J.* 17: 63–71.
- Escher, D., M. Bodmer-Glavas, A. Barberis, and W. Schaffner, 2000 Conservation of glutamine-rich transactivation function between yeast and humans. *Mol. Cell. Biol.* 20: 2774–2782.
- Fondon, J. W., E. A. D. Hammock, A. J. Hannan, and D. G. King, 2008 Simple sequence repeats: genetic modulators of brain function and behavior. *Trends Neurosci.* 31: 328–334.
- Gan, X., O. Stegle, J. Behr, J. G. Steffen, P. Drewe *et al.*, 2011 Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature* 477: 419–423.
- Gemayel, R., M. D. Vences, M. Legendre, and K. J. Verstrepen, 2010 Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu. Rev. Genet.* 44: 445–477.
- Hannan, A. J., 2010 Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for “missing heritability”. *Trends Genet.* 26: 59–65.
- Herrero, E., E. Kolmos, N. Bujdoso, Y. Yuan, M. Wang *et al.*, 2012 EARLY FLOWERING4 recruitment of EARLY FLOWERING3 in the nucleus sustains the *Arabidopsis* circadian clock. *Plant Cell* 24: 428–443.
- Hicks, K. A., A. J. Millar, I. A. Carre, D. E. Somers, M. Straume *et al.*, 1996 Conditional circadian dysfunction of the *Arabidopsis* early-flowering 3 mutant. *Science* 274: 790–792.
- Huang, H., S. Alvarez, R. Bindbeutel, Z. Shen, M. J. Naldrett *et al.*, 2016 Identification of evening complex associated proteins in *Arabidopsis* by affinity purification and mass spectrometry. *Mol. Cell. Proteomics* 15: 201–217.
- Jacob, F., 1977 Evolution and tinkering. *Science* 196: 1161–1166.
- Jain, R., M. C. Rivera, and J. A. Lake, 1999 Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl. Acad. Sci. USA* 96: 3801–3806.
- Jimenez-Gomez, J. M., J. A. Corwin, B. Joseph, J. N. Maloof, and D. J. Kliebenstein, 2011 Genomic analysis of QTLs and genes altering natural variation in stochastic noise. *PLoS Genet.* 7: e1002295.
- Kashi, Y., D. King, and M. Solter, 1997 Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.* 13: 74–78.
- Khattak, A. K., 2014 Natural variation in *Arabidopsis thaliana* growth in response to ambient temperatures. Ph.D. Thesis, University of East Anglia, Norwich, England.
- Kleinboelting, N., G. Huep, A. Kloetgen, P. Viehoveer, and B. Weisshaar, 2012 GABI-Kat SimpleSearch: new features of the *Arabidopsis thaliana* T-DNA mutant database. *Nucleic Acids Res.* 40: D1211–D1215.
- Kolmos, E., E. Herrero, N. Bujdoso, A. J. Millar, R. Tóth *et al.*, 2011 A reduced-function allele reveals that EARLY FLOWERING3 repressive action on the circadian clock is modulated by phytochrome signals in *Arabidopsis*. *Plant Cell Online* 23: 3230–3246.
- Kover, P. X., W. Valdar, J. Trakalo, N. Scarcelli, I. M. Ehrenreich *et al.*, 2009 A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* 5: e1000551.
- Lachowiec, J., C. Queitsch, and D. J. Kliebenstein, 2016 Molecular mechanisms governing differential robustness of development and environmental responses in plants. *Ann. Bot.* 117:795–809.
- Lander, E. S., and D. Botstein, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185–199.
- Lempe, J., J. Lachowiec, A. M. Sullivan, and C. Queitsch, 2013 Molecular mechanisms of robustness in plants. *Curr. Opin. Plant Biol.* 16: 62–69.
- Li H., 2013 Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv13033997 Q-Bio*.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079.
- Liu, X. L., M. F. Covington, C. Fankhauser, J. Chory, and D. R. Wagner, 2001 ELF3 encodes a circadian clock-regulated nuclear protein that functions in an *Arabidopsis* PHYB signal transduction pathway. *Plant Cell* 13: 1293–1304.
- Lu, X., and R. M. Murphy, 2015 Asparagine repeat peptides: aggregation kinetics and comparison with glutamine repeats. *Biochemistry* 54: 4784–4794.
- Möckli, N., and D. Auerbach, 2004 Quantitative β -galactosidase assay suitable for high-throughput applications in the yeast two-hybrid system. *Biotechniques* 36: 872–876.
- Nieto, C., V. López-Salmerón, J.-M. Davière, and S. Prat, 2014 ELF3-PIF4 interaction regulates plant growth independently of the evening complex. *Curr. Biol.* 25: 187–193.
- Nusinow, D. A., A. Helfer, E. E. Hamilton, J. J. King, T. Imaizumi *et al.*, 2011 The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* 475: 398–402.
- Orr, H. T., and H. Y. Zoghbi, 2007 Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* 30: 575–621.
- Perutz, M. F., T. Johnson, M. Suzuki, and J. T. Finch, 1994 Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc. Natl. Acad. Sci. USA* 91: 5355–5358.
- Press, M. O., K. D. Carlson, and C. Queitsch, 2014 The overdue promise of short tandem repeat variation for heritability. *Trends Genet.* 30: 504–512.
- Press, M. O., A. Lanctot, and C. Queitsch, 2016 PIF4 and ELF3 act independently in *Arabidopsis thaliana* thermoresponsive flowering. *PLoS One* 11: e0161791.
- Queitsch, C., K. D. Carlson, and S. Girirajan, 2012 Lessons from model organisms: phenotypic robustness and missing heritability in complex disease. *PLoS Genet.* 8: e1003041.
- R Development Core Team, 2016 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rowan, B. A., V. Patel, D. Weigel, and K. Schneeberger, 2015 Rapid and inexpensive whole-genome genotyping-by-sequencing for

- crossover localization and fine-scale genetic mapping. *G3* 5: 385–398.
- Rutherford, S. L., and S. Lindquist, 1998 Hsp90 as a capacitor for morphological evolution. *Nature* 396: 336–342.
- Sangster, T. A., A. Bahrami, A. Wilczek, E. Watanabe, K. Schellenberg *et al.*, 2007 Phenotypic diversity and altered environmental plasticity in *Arabidopsis thaliana* with reduced Hsp90 levels. *PLoS One* 2: e648.
- Sangster, T. A., N. Salathia, H. N. Lee, E. Watanabe, K. Schellenberg *et al.*, 2008a HSP90-buffered genetic variation is common in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 105: 2969–2974.
- Sangster, T. A., N. Salathia, S. Undurraga, R. Milo, K. Schellenberg *et al.*, 2008b HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. *Proc. Natl. Acad. Sci. USA* 105: 2963–2968.
- Schaefer, M. H., E. E. Wanker, and M. A. Andrade-Navarro, 2012 Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks. *Nucleic Acids Res.* 40: 4273–4287.
- Sharma, D., S. Sharma, S. Pasha, and S. K. Brahmachari, 1999 Peptide models for inherited neurodegenerative disorders: conformation and aggregation properties of long polyglutamine peptides with and without interruptions. *FEBS Lett.* 456: 181–185.
- Stott, K., J. M. Blackburn, P. J. Butler, and M. Perutz, 1995 Incorporation of glutamine repeats makes protein oligomerize: implications for neurodegenerative diseases. *Proc. Natl. Acad. Sci. USA* 92: 6509–6513.
- Sureshkumar, S., M. Todesco, K. Schneeberger, R. Harilal, S. Balasubramanian *et al.*, 2009 A genetic defect caused by a triplet repeat expansion in *Arabidopsis thaliana*. *Science* 323: 1060–1063.
- Szamecz, B., G. Boross, D. Kalapis, K. Kovács, G. Fekete *et al.*, 2014 The genomic landscape of compensatory evolution. *PLoS Biol.* 12: e1001935.
- Taipale, M., D. F. Jarosz, and S. Lindquist, 2010 HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* 11: 515–528.
- Tajima, T., A. Oda, M. Nakagawa, H. Kamada, and T. Mizoguchi, 2007 Natural variation of polyglutamine repeats of a circadian clock gene *ELF3* in *Arabidopsis*. *Plant Biotechnol.* 24: 237–240.
- Undurraga, S. F., M. O. Press, M. Legendre, N. Bujdosó, J. Bale *et al.*, 2012 Background-dependent effects of polyglutamine variation in the *Arabidopsis thaliana* gene *ELF3*. *Proc. Natl. Acad. Sci. USA* 109: 19363–19367.
- Usdin, K., 2008 The biological effects of simple tandem repeats: lessons from the repeat expansion diseases. *Genome Res.* 18: 1011–1019.
- Wang, Y., J. Lu, J. Yu, R. A. Gibbs, and F. Yu, 2013 An integrative variant analysis pipeline for accurate genotype/haplotype inference in population NGS data. *Genome Res.* 23: 833–842.
- Winter, D., B. Vinegar, H. Nahal, R. Ammar, G. V. Wilson *et al.*, 2007 An “electronic fluorescent pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2: e718.
- Yoshida, R., R. Fekih, S. Fujiwara, A. Oda, K. Miyata *et al.*, 2009 Possible role of early flowering 3 (*ELF3*) in clock-dependent floral regulation by short vegetative phase (SVP) in *Arabidopsis thaliana*. *New Phytol.* 182: 838–850.
- Yu, J.-W., V. Rubio, N.-Y. Lee, S. Bai, S.-Y. Lee *et al.*, 2008 COP1 and *ELF3* control circadian function and photoperiodic flowering by regulating GI stability. *Mol. Cell* 32: 617–630.
- Zagotta, M. T., K. A. Hicks, C. I. Jacobs, J. C. Young, R. P. Hangarter *et al.*, 1996 The *Arabidopsis* *ELF3* gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. *Plant J.* 10: 691–702.
- Zhao, L., M. Nakazawa, T. Takase, K. Manabe, M. Kobayashi *et al.*, 2004 Overexpression of *LSH1*, a member of an uncharacterised gene family, causes enhanced light regulation of seedling development. *Plant J.* 37: 694–706.
- Zhao, R., M. Davey, Y.-C. Hsu, P. Kaplanek, A. Tong *et al.*, 2005 Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell* 120: 715–727.
- 1001 Genomes Consortium, 2016 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* 166: 481–491.

Communicating editor: K. Bomblies

SUPPLEMENTARY INFORMATION – PRESS AND QUEITSCH 2016

SUPPLEMENTARY TEXT (File S1)

Ws genome data. We resequenced the genome of the Ws (Wassilewskija) parent to ~12X coverage, calling 3.2×10^5 variants relative to Col, and used these variants as identifiers for Ws haplotype blocks. Comparisons of single nucleotide variants (SNVs) with related *A. thaliana* strains Ws-2 and Ws-0 yielded substantial overlap but also significant variation (Figure S2), indicating good SNV ascertainment. For each F_2 individual, we mapped reads to the Col-0 reference and imputed haplotypes across each of the 96 individuals to identify chromosome blocks inherited from each parent (Figure S3).

Resequencing data processing. Where appropriate, FASTQ files were demultiplexed according to barcodes (Table S3) using the `illumina_demultiplexer.py` script from the University of Idaho (https://github.com/ibest/GRC_Scripts/blob/master/Illumina_demultiplexer.py). Following alignment, duplicate removal, BAM conversion, and indexing, the following command was used to generate VCF files:

```
$ samtools mpileup -uf TAIR10.fas.masked [bam file list] bcftools view -vcg - > [output.vcf]
```

Where TAIR10.fas.masked is the repeat-masked TAIR10 reference genome build in FASTA format. For SNVs tagging differences between Col and Ws, SNVs from F_2 s were filtered for membership in the Ws SNV call set, >1.5X average depth across libraries, and overall allele frequency 0.3-0.7 across libraries. This yielded approximately 230,000 variants. VCFs were then split by chromosome for SNPtools analysis. For each chromosome, commands were run according to SNPtools documentation

(https://www.hgsc.bcm.edu/sites/default/files/software/SNPTools/V1_0/SNPTools-documentation-v1.1.pdf).

SNPtools imputation was used to assign genotypes across this large set of variants to the various F_2 s, resulting in a final VCF file with genotype probabilities across each. This file was parsed into a simple genotype table, accepting the most probable genotype in each case. This table was used for QTL analysis and updated with selected PCR genotyping loci as indicated. Both VCF file and genotype table (in CSV format for QTL analysis) are provided at <https://figshare.com/s/e01a40b98a4ef5a9e5b3>.

QTL analysis: These genotypes were used for QTL analysis mapping the genetic basis of variation in hypocotyl length of segregants. Twelve loci were found to have strongly discordant genotypes (requiring many implausible recombinations), and were discarded before map re-estimation. Additional individuals were genotyped by PCR at informative markers and added to the analysis. One- and two-dimensional QTL analyses were performed using the commands `scanone()` and `scantwo()` using imputation to estimate QTLs. Notably, it was necessary to include phenotypes of all phenotyped F_2 s (not just the extremes), with genotypes marked as missing, to avoid effect size overestimation.

Interval mapping was performed with the *bayesint()* function. Multiple QTL mapping was performed with the *mapqtl()* function and refinement with the *refineqtl()* function.

The direct test for epistasis with *ELF3* was implemented in the following fashion. We computed the difference in LOD profiles across all loci between segregants homozygous for the Col vs. Ws alleles of *ELF3* (effectively, the difference between the two curves in Figure 2B). This “delta LOD” profile captures the difference in apparent contributions of loci between segregants homozygous for each allele of *ELF3*. However, there is no obvious null distribution for this difference in profiles. Previously, this has been addressed by refitting QTLs to permuted phenotype data to estimate parameters of this null distribution (Sangster *et al.* 2008). We chose a similar approach, by drawing random samples of genotyped individuals of equal size to the *ELF3*-Col/*ELF3*-Col and *ELF3*-Ws/*ELF3*-Ws samples. For 1000 such pairs of random samples, we performed QTL scans for each sample, and drew the null distribution from the difference between each such pair of scans. In contrast with the previous method, we elected to more conservatively use the permutation QTL differences directly to define an empirical nonparametric null distribution. From this distribution we set a 5% confidence threshold across the entire profile (Figure 2C).

Y2H screen: pGBKT7-*ELF3*-7Q, pGBKT7-*ELF3*-16Q, pGBKT7-*ELF3*-0Q, and pGBKT7-*ELF3*-23Q were transformed into yeast strain AH109, and pGADT7-*ELF4* was transformed into yeast strain Y187 (Gietz and Woods 2006), along with their respective empty vectors. These transformants were mated, and positive yeast two-hybrid (Y2H) interaction was confirmed on C-leu-trp-his (-LTH) plates incubated for 4-7d at 23° before continuing with the Y2H screen. >200 positive colonies from the screen were patched to C-LTH and C-Ade to confirm reporter activation, and true activating colonies were analyzed by insert PCR and sequencing. Transformants for each clone were mated to AH109 carrying pGBKT7 as an empty vector or with various *ELF3* inserts to test *ELF3*-prey interactions. We then tested each clone for independent activation of the *ADE2* reporter, which reduced this pool to 67 clones (tabulated in detail in File S2). Removing duplicates and clones carrying multiple plasmids, these represented 36 unique coding DNA sequences. Four clones expressed cDNAs corresponding to known *ELF3* interactor *ELF4*. Prey plasmid inserts were isolated from yeast using the Zymo Research Yeast Plasmid Prep kit (Irvine, CA) and re-transformed into the Y187 strain. Clones showing positive activation with *ELF3*-7Q were further transformed into chemically competent TOP10 *E. coli*, prepared with the ProMega plasmid prep kit (Fitchburg, WI), and re-tested in Y187 and PJ69-4α. We were able to confirm interactions for 3 of these clones.

Each of these clones thus went through 5 rounds of confirmation. Results shown are from PJ69-4α as MATα parent, which showed more robust activation (grew on C-LTHA, which Y187 did not).

SUPPLEMENTARY FIGURES

Figure S1. Workflow for genetic analysis of the Col x Ws F₂s. Seedling plate image indicates hypocotyl length phenotype variation among Col x Ws segregants in red. Histogram shows phenotype distribution for Col x Ws segregants. Gel indicates different sizes of the *ELF3* STR between Col and Ws and in F₂s.

Figure S2. Overlap of high-confidence ($Q \geq 40$) single nucleotide variants (SNVs) observed in the Ws strain with two related strains.

Figure S3. Genotypes of the 96 extreme F_2 s used for sequencing. Each panel sorted vertically from longest to shortest hypocotyls, chromosomes and markers are displayed along the X axis. *ELF3-Col* homozygotes at top, *ELF3-Ws* homozygotes at bottom. Red indicates homozygosity for Col allele, blue indicates heterozygosity, and green indicates homozygosity for Ws allele. *ELF3* locus is indicated.

Figure S4. 2-locus QTL scan indicates potential genetic interactions between *ELF3* and other loci. Lower triangle indicates LOD for full model, upper triangle indicates LOD for interaction terms alone. *ELF3* corresponds to the locus on chromosome 2.

Figure S5. Effect on hypocotyl length of various QTLs interacting with the *ELF3* locus. X-axis is non-*ELF3* locus, different lines indicate effects of other locus in different *ELF3* backgrounds. Loci are taken from the midpoint estimates of the refined QTL model (in which all QTLs are jointly considered). Error bars represent standard error of the mean.

Figure S6. Analysis of 15d hypocotyl length under SD of candidate gene mutants.

N ~ 50 for each line, each panel represents an independently conducted experiment, each repeated at least once with similar results. (A): T-tests for differences vs. Col WT: $p < 2.2e-16$ for *pif4* and *elf3*; $p = 0.009$ for *cyct1;4*, $p = 0.033$ for *nup98*, $p > 0.1$ for all others. The *cyct1;4* T-DNA insertion could not be confirmed by PCR and was thus omitted from further analysis. (B): Same as (A), but testing other candidates. *lsh9* is ~0.5 mm shorter ($p = 0.03$, t-test), *elf3* is very elongated relative to WT ($p < 2.2e-16$, t-test). All others $p > 0.1$. N > 45 for each genotype. Experiment was repeated with similar results. (C): Double mutant analysis of *nup98* and *elf3* does not indicate a genetic interaction. Error bars indicate standard deviation; N > 40 for each group; F-statistic and p-value are displayed for testing the hypothesis of an interaction between the two genes in controlling phenotype.

Figure S7. GAL4 DBD-ELF3 fusions do not lead to Y2H activation in combination with GAL4 AD-LSH9 vectors carrying either the Col or Ws variant of *LSH9*. Yeast carrying indicated vectors were spotted in five-fold dilutions onto C-leu-trp (C-LT) or C-leu-trp-his-ade (C-LTHA) media. For each protein X, experiments were repeated with independent PJ69-4α + pGADT7-X transformants with similar results.

Figure S8. *ELF3* expression is not substantially affected by *Ish9* T-DNA insertion. *ELF3* expression was measured by PCR and normalized to *UBC21* expression (Col set to 1.0). Error bars are standard error of the mean from three technical replicates. This experiment was repeated with similar results.

Figure S9. The GAL4 DBD-ELF3 fusion does not autoactivate in combination with a GAL4 AD-empty vector (pGADT7). Yeast carrying indicated vectors were spotted in five-fold dilutions onto C-leu-trp (C-LT) or C-leu-trp-his-ade (C-LTHA) media. This experiment was repeated with an independent PJ69-4 α + pGADT7 transformant with similar results.

Figure S10. ELF3-PIF4 interaction not detectable in our assay. Row labels of strains correspond to contents of pGBK vector expressing GAL4 DBD fusions. Label at right indicates GAL4 AD fusion expressed in strains. (A): Assays in PJ69-4 α strain background, (B): Assays in Y187 strain background. C-LT: YC-leu-trp media, C-LTH: YC-leu-trp-his media, YC-leu-trp-his-ade media. Y187 background showed activation with BD-ELF3 and AD-ELF4 and AD-PHYB (data not shown), but only on C-LTH. For each protein, experiments were repeated with independent PJ69-4 α /Y187 + pGADT7-PIF4 transformants with similar results.

SUPPLEMENTARY TABLES

Table S1. Primers employed in this study.

primer name	target	sequence	note
nga111_F	chr1 mapping	TGTTTTTTAGGACAAATGGCG	
nga111_R	chr1 mapping	CTCCAGTTGGAAGCTAAAGGG	
nga280_F	chr1 mapping	GGCTCCATAAAAAGTGCACC	
nga280_R	chr1 mapping	CTGATCTCACGGACAATAGTGC	
nga168_f	chr2 mapping	GAGGACATGTATAGGAGCCTCG	
nga168_R	chr2 mapping	TCGTCTACTGCACTGCCG	
1H1L-1.6_F	chr4 mapping	CTAGAGCTTGAAAGTTGATG	
1H1L-1.6_R	chr4 mapping	TTGAGTCCTTCTTGCTG	
g3883-1.4_F	chr4 mapping	TGTTTCAGAGTAGCCAATTC	
g3883-1.4_R	chr4 mapping	CATCCATCAAACAACTCC	
nga129_F	chr5 mapping	CACACTGAAGATGGTCTTGAGG	
nga129_R	chr5 mapping	TCAGGAGGAACTAAAGTGAGGG	
rps4_NT_R	chr5 mapping	GGTACCTTTTCTTGCAATG	
rps4-NT_F	chr5 mapping	TCATCAATTTCCGTGGGGCA	
cyct1;4 tdna confirm F	cyct1;4	CTTTCCATTTTCTCTTCCGG	failed
cyct1;4 tdna confirm R	cyct1;4	TGGAATTTTGTTAACGATGGG	failed
ELF3 STR F (w ncoi site)	ELF3	CCCTTC CCATGG GATTCCACCTCCTGGTAAT	
ELF3 STR R (w nari site)	ELF3	TTTTGGGGCGCC GGTTCAATTATCAGATCTCTG	
ELF3 y2h cloning F	ELF3	AG GCC GAA TTC CCG GGG ATG AAG AGA GGG AAA GA	
ELF3 y2h cloning R	ELF3	G CAG GTC GAC GGATCC TTA AGG CTT AGA GGA GTC ATA GCG	
ELF4 y2h cloning f	ELF4	G GAG GCC AGT GAATTC ATG AAG AGG AAC GGC GAG ACG AAA	
ELF4 y2h cloning r	ELF4	C ATC TGC AG CAT ATG C TTA AGC TCT AGT TCC GGC AGC ACC	
lsh9 y2h clone f	lsh9	G GAG GCC AGT GAATTC ATGTCTTCGGATCGTCACACACC	
lsh9 y2h clone r	lsh9	C ATC TGC AG C TCG AGC TCAAGTGAACGACTGATTAGGAG	
lsh9_expr_f	LSH9	CGGGTTCACCTTGAGGGAAGTG	use with R lsh9 genotyping primer
lsh9-1 genotyping_r	LSH9	CAAGTGAACGACTGATTAGGAGT	
lsh9-1 genotyping_f	LSH9	TTTTCGTGAATAATTCTCTCTCCC	
nup98 tdna confirm f	nup98	TTTTCTGAATTTTGGCTTTGTCTT	
nup98 tdna confirm r	nup98	AGTCTACACACATCCGTCTCTCCT/AGTCTACACACATCCGTCTCT	
matchmaker insert amp F	pGADT7/ pGBKT7 plasmid insert	TCA TCG GAA GAG AGT AG	
matchmaker insert amp R	pGADT7/ pGBKT7 plasmid insert	GTCACCTTTAAATTTGTATAC	

phyb_ cterm clone f	y2h	phyb	GGAGGCCAGTCATATGATTGATGAGTTAGGTGCAGTTG	
phyb_ cterm clone r	y2h	phyb	CATCTGCAGGGATCCCCTAATATGGCATCATCAGCATCAT	
pif4 y2h cloning f		pif4	GGAGGCCAGTGAATTCATGGAACACCAAGGTTGGAGTTTT	
pif4 y2h cloning r		pif4	C ATC TGC AG C TCG AGC CTAGTGGTCCAAACGAGAACCGTC	
pif4-2 genotyping f		pif4	ACCTCCTCAAGTCATGGTTAAGCCTAAGCC	
pif4-2 genotyping r		pif4	TCCAAACGAGAACCGTCGGT	
saur38 tdna confirm f		saur38	GATGTTAGATCAATGGGCTG	
saur38 tdna confirm r		saur38	CCCCGATTCTCAATATTCACC	
soc1 tdna confirm f		soc1	AAGGATGAGGTTTCAAGCGTC	
soc1 tdna confirm r		soc1	TGGCGAATTCATAAAGTTTGC	
svl2-1 genotyping		svl2	GAGTGTGCATTTGTGTGAAG	
svl2-1 genotyping		svl2	ACGATGCGTTCTATGAGCAA	
pif4-2 T-DNA primer		T-DNA	TAGCATCTGAATTTTCATAACCAATCTCGATACAC	
"o8474", lsh9-1 genotyping_tdna		T-DNA (GABI-KAT)	ATAATAACGCTGCGGACATCTACATTTT	
tsf-1 genotyping lp		TSF	CGGTAACTTGATTTTGTTC	
tsf-1 genotyping rp		TSF	ACGTGGACTCTCGTAGCACA	
UBC21 expression f		UBC21	GACCAAGATATTCATCCTA	
UBC21 expression r		UBC21	GTAAAGAGGACTGTCCG	

Table S2. Mutant lines employed in this study.

mutant	Identifier/origin
soc1 (Col-0):	SALK_138131C
<i>nup98</i> (Col-0):	CS2103568 = GK-288A08-015346
arf1-3 (Col-0):	CS24599
pif4-2 (Col-0):	CS66043
<i>saur38</i> (Col-0):	SALK_001155
cyct1;4 (Col-0):	SALK_139322
<i>lsh9</i> (Col-0)	CS328437
<i>tsf</i> (Col-0)	SALK_087522C
<i>elf3-200</i> (Col)	GABI750E02
<i>elf3-4</i> (Ws)	CS3790
<i>svl2</i> (Col)	SALK_057865
<i>elf3-200 nup98</i>	This study
<i>elf3-200 lsh9</i>	This study

Table S3. Col x Ws segregants analyzed by genotyping-by-sequencing.

Individual	barcode	reads mapped	genome coverage	growth plate number	plate position	d15 hypocotyl length	transplant tray	row (A-F)	column (1-12)	ELF3 Col alleles
A1	GACTTC GTA	2295328	1.926477 847	5	1	5.669	BOTTOM1	A	7	0
A2	AATCTC GTC	2558578	2.147424 611	5	2	6.127	BOTTOM1	A	8	0
A3	TCTTAC GGT	3331821	2.796410 512	5	4	11.236	TOP1	B	2	2
A4	TTGCCA CTA	2354351	1.976016 084	5	5	9.591	TOP1	B	3	0
A5	GCGTTA ATG	3567044	2.993834 104	5	10	11.24	TOP1	B	4	2
A6	CTTCAA CGA	1349755	1.132854 697	6	9	7.457	BOTTOM4	A	6	2
A7	AGCGTA CTA	2054778	1.724583 283	7	5	11.136	TOP1	B	7	2
A8	TGATCT GAG	6217132	5.218063 419	8	4	9.95	TOP1	B	8	0
A9	TACGGT GAA	2410759	2.023359 541	10	6	9.611	TOP1	B	11	0
A10	ATGCAT GAC	2645891	2.220706 756	10	9	9.539	TOP1	B	12	0
A11	AACTGT CCT	2747651	2.306114 326	11	3	6.543	BOTTOM1	B	11	0
A12	AGACAC CAT	3354860	2.815747 235	11	5	6.15	BOTTOM1	B	12	2
B1	AGAACA TCC	1881398	1.579064 765	14	5	6.347	BOTTOM1	C	8	0
B2	CTGGAT AAG	2011360	1.688142 384	16	4	6.33	BOTTOM1	D	3	0
B3	GTGACT TAC	3836304	3.219825 084	17	7	6.014	BOTTOM1	D	6	0
B4	CTCACT GAA	1059679	0.889392 766	18	7	10.985	TOP1	C	9	2
B6	TGAGAC ATG	2240186	1.880196 949	19	2	7.147	BOTTOM1	D	10	2
B7	TCCACG ATT	3525873	2.959279 121	19	3	11.082	TOP1	C	10	2
B8	GACTGA TAC	1643219	1.379160 19	20	1	6.238	BOTTOM1	D	11	0
B9	ACATCT GCT	2702340	2.268084 625	22	5	11.093	TOP1	D	2	2
B10	ACGTTA GGA	3353328	2.814461 422	25	1	6.038	BOTTOM1	E	9	0
B11	TGGACT GTT	3052538	2.562007 188	25	2	7.366	BOTTOM1	E	10	2
B12	TTCTCA AGG	4092160	3.434566 035	26	2	10.628	TOP1	D	9	0
C1	CAGGCT TAA	2376433	1.994549 594	26	8	5.687	BOTTOM1	E	11	0
C2	GTGGTC TAA	1971626	1.654793 481	27	5	6.623	BOTTOM1	F	2	2
C3	CACTAG ACA	1714651	1.439113 35	27	7	6.099	BOTTOM1	F	3	0
C4	CTCGAT CTA	2346675	1.969573 587	27	8	5.841	BOTTOM1	F	4	0
C5	TGGCAT TCA	2924555	2.454590 551	28	8	7.356	BOTTOM1	F	7	2
C6	ACATTG CAC	159016	0.133462 756	29	10	10.164	TOP1	E	3	0
C7	TATGCC ACT	1974750	1.657415 467	30	9	9.948	TOP1	E	7	0

C8	ACCGAT CAT	1621564	1.360985 064	34	4	6.036	BOTTOM2	A	2	0
C9	AGGTA GACT	3366856	2.825815 526	34	8	7.093	BOTTOM2	A	3	2
C10	CAGTTG ATC	3891209	3.265907 068	35	8	13.69	TOP1	E	11	2
C11	ACTACC GAT	394870	0.331415 949	36	7	6.536	BOTTOM2	A	8	0
C12	CATCAT ACG	455273	0.382112 425	36	8	6.947	BOTTOM2	A	9	2
D1	CGTCAG TAA	2644128	2.219227 064	37	5	6.455	BOTTOM2	A	10	0
D2	TTACGA CAG	4586816	3.849732 767	37	6	11.178	TOP1	F	6	2
D3	ACCACT TGA	2617445	2.196831 916	38	2	11.28	TOP1	F	7	2
D4	GAGATC GAT	2974369	2.496399 638	38	5	10.384	TOP1	F	8	0
D5	CATAGT TGC	1776793	1.491269 376	39	8	9.551	TOP1	F	11	0
D6	GAAGTC ATC	1235929	1.037320 086	41	3	7.123	BOTTOM2	B	4	2
D7	TGTAAC GTC	410208	0.344289 193	42	3	9.333	TOP4	D	12	0
D8	TAGCAG ATG	618152	0.518817 413	43	2	11.533	TOP2	A	5	2
D9	CCGACT AAT	2249982	1.888418 771	44	4	7.572	BOTTOM4	E	7	2
D10	AGTACG CTA	6326180	5.309587 836	47	5	7.401	BOTTOM4	F	1	2
D11	ATTGCT TGG	1483518	1.245122 511	52	3	11.307	TOP2	B	10	2
D12	CAATTA CGG	3716441	3.119223 596	52	6	11.585	TOP2	B	11	0
E1	GGCAAT TCT	3406103	2.858755 688	53	9	11.33	TOP2	C	4	2
E2	GTTACA GGT	2893975	2.428924 636	54	4	7.08	BOTTOM2	C	11	2
E3	ACACTG ATG	1969162	1.652725 436	54	10	7.558	BOTTOM4	F	11	2
E4	TGAGG ATTC	74043	0.062144 582	55	3	11.25	TOP2	C	7	2
E5	GTATAC AGC	1951625	1.638006 563	55	6	11.276	TOP2	C	8	2
E6	ATAGCC TTC	1523251	1.278470 575	56	9	13.121	TOP2	C	11	2
E7	ATCGAT TCC	360569	0.302626 984	57	8	11.378	TOP2	D	2	0
E8	ACCATA AGC	1291511	1.083970 278	60	10	11.106	TOP2	D	8	2
E9	AATGCG TCA	2193069	1.840651 465	62	3	5.465	BOTTOM2	D	9	2
E10	AGAATC GTG	4479066	3.759297 767	63	1	11.802	TOP2	D	12	2
E11	GAGGC AATA	6544234	5.492601 418	63	2	6.727	BOTTOM2	D	10	2
E12	CTCTGA CTT	2837854	2.381822 06	63	7	7.325	BOTTOM2	D	12	2
F1	CTCAGC TAT	2162856	1.815293 575	64	10	7.492	BOTTOM5	B	1	2
F2	TTGTGA GCA	803296	0.674209 502	66	1	11.79	TOP2	E	6	2
F3	TCAGAA GAC	3080338	2.585339 838	66	8	9.554	TOP2	E	10	0

F4	AATCGG CAA	3392517	2.847352 904	68	1	6.247	BOTTOM2	F	1	0
F5	GTATTC GAG	2432555	2.041653 01	71	2	7.447	BOTTOM5	B	5	2
F6	TAGAAG CCA	1147423	0.963036 651	72	2	5.903	BOTTOM2	F	10	0
F7	AATGGC CTT	74709	0.062703 558	72	4	7.5	BOTTOM5	B	7	2
F8	GGTCAA CTT	1528117	1.282554 628	72	7	9.469	TOP5	B	7	0
F9	TCAGTT CAG	1969461	1.652976 388	74	6	6.274	BOTTOM3	A	1	0
F10	TACACC TAG	4598401	3.859456 104	75	5	11.728	TOP3	A	3	2
F11	CTAAGA TGG	2896390	2.430951 555	76	7	6.275	BOTTOM3	A	4	0
F12	ATGTCG TAG	3236556	2.716454 221	79	1	11.183	TOP3	A	11	0
G1	CATGTT GGA	2819706	2.366590 372	83	10	10.927	TOP3	B	12	2
G2	GATGAT AGC	1600038	1.342918 207	84	10	7.129	BOTTOM3	B	6	2
G3	TGAAGC TGT	487191	0.408901 329	85	3	9.945	TOP3	C	3	0
G4	GCAATG ACT	2869408	2.408305 456	86	7	7.122	BOTTOM3	B	8	2
G5	TCGAGA TCT	80032	0.067171 173	88	1	6.308	BOTTOM3	B	10	2
G6	CAAGA GCAT	4776085	4.008586 986	89	6	9.568	TOP3	C	8	0
G7	ACGCAC TAA	2346562	1.969478 746	91	6	11.414	TOP3	C	9	2
G8	GAACTA GGA	1335094	1.120549 662	92	9	10.626	TOP3	D	2	0
G9	TAATGC ACC	2070382	1.737679 782	92	10	9.545	TOP3	D	3	0
G10	ATCCTC AGT	572433	0.480445 276	93	5	9.986	TOP3	D	4	0
G11	TGACTG CTA	3446213	2.892420 169	93	8	9.61	TOP3	D	5	0
G12	TTCGAG TGA	677494	0.568623 387	94	4	12.438	TOP3	D	8	2
H1	CTGATT GGT	2572313	2.158952 451	94	5	9.509	TOP3	D	9	0
H2	GTCCAT ATC	2809401	2.357941 345	94	7	5.843	BOTTOM3	C	6	0
H3	GATCTC TCA	1898536	1.593448 756	95	1	6.764	BOTTOM3	C	7	2
H4	CTAATG GCA	165936	0.139270 74	103	5	6.028	BOTTOM3	D	8	0
H5	TGCGTT AGT	1345594	1.129362 354	103	7	9.381	TOP5	F	1	0
H7	ATATCG ACC	1170512	0.982415 34	105	3	13.88	TOP3	E	12	2
H8	ATAACT CGC	1359083	1.140683 725	107	7	6.881	BOTTOM3	E	7	2
H9	TACTCA GTC	4314258	3.620973 762	108	5	6.35	BOTTOM3	E	9	0
H10	GGAGT ATCT	2714867	2.278598 585	109	1	6.465	BOTTOM3	E	10	0
H11	AACTCT AGC	3458587	2.902805 716	109	10	9.839	TOP3	F	9	0
H12	CGCTAC ATT	3893747	3.268037 221	110	5	10.838	TOP3	F	10	2

B5	CATTAC GAG	2578299	2.163976 524	111	1	5.796	BOTTOM3	F	4	0
H6	GTGCTA CAT	1661072	1.394144 284	113	2	6.032	BOTTOM3	F	12	0
	Mean	2374934. 3	1.993291 757							

Table S4. Multiple QTL mapping reveals principal effects of *ELF3* locus.

Chromosome	Estimated Position (bp)	90% credible interval (positions in bp)	LOD
<i>Raw model</i>			
1	21,727,855	(21,177,007 : 25,669,040)	4.29
2	6,833,671 ¹	(2,331,734 : 19,411,531) ¹	2.22
4	10,602,975	(2,823,163 : 18,520,802)	2.58
5	18,703,510	(8,622,189 : 26,967,913)	2.88
<i>Refined model (fitting all QTLs)</i>			
1	25,906,082	(4,356,888 : 30,407,116)	2.92
2	10,906,049 ¹	(10803445 : 11446925) ¹	14.98
2	19,411,531	(1,446,925 : 19,411,531)	2.03
4	10,661,954	(5,631,553 : 12,314,751)	4.36
5	6,142,833	(3,073,579 : 26,967,913)	3.11

¹: *ELF3* is tagged by a polymorphism at position 11,063,324 on chromosome 2.

File S2. Clones recovered as Y2H interactors of ELF3. First sheet of spreadsheet ("clones") lists and tabulates properties of all clones isolated from a Y2H screen of the ELF3 protein against the *A. thaliana* proteome, including BLAST hits and redundancy. Second sheet ("metadata") explains fields of first sheet in detail. (.xlsx, 74 KB)

www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.193359/-/DC1/FileS2.xlsx