

1 *The variable ELF3 polyglutamine tract mediates complex epistatic interactions in*
2 *Arabidopsis thaliana.*

3

4 Maximilian Oliver Press, Christine Queitsch

5

6 University of Washington Department of Genome Sciences

7 Seattle, WA 98195

8 Correspondence to queitsch@uw.edu

9

1 Running title: The ELF3 polyQ is an epistatic hub.

2

3 Keywords: ELF3, polyglutamine, short tandem repeat, microsatellite, epistasis,
4 robustness

5

6 Correspondence to:

7 Christine Queitsch

8 Foege S-410

9 Box 355065

10 Foege Building, Room S410B

11 3720 15th Ave NE

12 University of Washington

13 Seattle, WA 98195-5065

14 206-685-8935

15 queitsch@uw.edu

16

ABSTRACT

1 Short tandem repeats are hypervariable genetic elements that occur frequently in
2 coding regions. Their high mutation rate readily generates genetic variation contributing
3 to adaptive evolution and human diseases. We recently proposed that short tandem
4 repeats are likely to engage in epistasis because they are well-positioned to
5 compensate for genetic variation arising at other loci due to their high mutation rate. We
6 previously reported that natural *ELF3* polyglutamine variants cause reciprocal genetic
7 incompatibilities in two divergent *Arabidopsis thaliana* backgrounds. Here, we dissected
8 the genetic architecture of this incompatibility and used a yeast two-hybrid strategy to
9 identify proteins whose physical interactions with *ELF3* were modulated by
10 polyglutamine tract length. Using these two orthogonal approaches, we identify specific
11 genetic interactions and physical mechanisms by which the *ELF3* polyglutamine tract
12 may mediate the observed genetic incompatibilities. Our work elucidates how short
13 tandem repeat variation, which is generally underascertained in population-scale
14 sequencing, can contribute to phenotypic variation. Furthermore, our results support our
15 proposal that highly variable STR loci can contribute disproportionately to the epistatic
16 component of heritability.
17

INTRODUCTION

18
19
20 Evolution is a tinkerer rather than a designer (Jacob 1977; Alon 2003); that is,
21 adaptations are generally short-term, incremental fixes rather than alterations in
22 fundamental biological plans. This principle is believed to underlie many design
23 properties of biological systems. Thus, many (or most) genetic adaptations may be
24 compensations for other genetic variants in a given background (Szamecz *et al.* 2014).
25 One abundant source of genetic variation for such tinkering lies in short tandem repeats
26 (STRs), genetic elements with high mutation rates. Due to these high mutation rates,
27 STRs may be more likely than substitutions to contribute adaptive variants on a per-
28 locus basis (Kashi *et al.* 1997; Gemayel *et al.* 2010; Hannan 2010). If ‘tinkering’ is a
29 dominant mode of adaptation, STRs should be likely to show epistasis with other loci,
30 and indeed, this expectation is borne out in the handful of well-characterized STRs
31 (Press *et al.* 2014).

32 One such STR resides in the *Arabidopsis thaliana* gene *ELF3*, where it encodes
33 a polyglutamine tract that varies in length across different natural strains (Tajima *et al.*
34 2007; Undurraga *et al.* 2012). We have previously shown that these *ELF3*-STR variants
35 have strong effects on phenotype, and that these effects differ depending on the genetic
36 background expressing a particular variant (Undurraga *et al.* 2012). These observations
37 suggest that background-specific variants are modifying the effect of STR alleles
38 through epistasis. The high variability of the *ELF3*-STR relative to expectations
39 suggests that this STR may compensate for many background-specific polymorphisms
40 across globally-distributed strains of *A. thaliana*.

41 *ELF3* has been previously identified as a plausible candidate gene underlying a
42 QTL for trait variance (*i.e.* noise) in the phenotypes of *A. thaliana* recombinant inbred
43 lines (Jimenez-Gomez *et al.* 2011; Lachowiec *et al.* 2015). These results invite
44 comparison to known ‘robustness genes’ such as HSP90 (Sangster *et al.* 2007, 2008a;

1 b), which can reveal or conceal the phenotypic consequences of many other genetic
2 variants. A mechanistic explanation of this robustness phenomenon is epistasis, in
3 which a robustness gene interacts with many other loci (Queitsch *et al.* 2012;
4 Lachowiec *et al.* 2015), as for the promiscuous chaperone HSP90 (Taipale *et al.* 2010).
5 Our previous findings and the many studies describing ELF3's crucial functions in plant
6 development lead us to hypothesize that *ELF3* lies at the center of an epistatic network
7 and that the ELF3's polyglutamine tract modifies these interactions.

8 It is well-established that ELF3 functions promiscuously as an adaptor protein in
9 multiple protein complexes that are involved in a variety of developmental pathways (Liu
10 *et al.* 2001; Yu *et al.* 2008; Yoshida *et al.* 2009; Nusinow *et al.* 2011; Chow *et al.* 2012).
11 Polyglutamine tracts such as the one encoded by the *ELF3*-STR often mediate protein
12 interactions (Perutz *et al.* 1994; Stott *et al.* 1995; Schaefer *et al.* 2012). Therefore, it is
13 plausible to assume that variation in the ELF3 polyglutamine tract affects ELF3's
14 interactions with its partner proteins. The ELF3 C-terminus, which contains the STR-
15 encoded polyglutamine tract, is necessary for nuclear localization (Herrero *et al.* 2012)
16 and ELF3 homodimerization (Liu *et al.* 2001), but thus far only one other protein
17 (Phytochrome Interacting Factor 4, PIF4) has been shown to interact with this ELF3
18 domain (Nieto *et al.* 2014). Thus, the phenotypic and epistatic effects of ELF3-polyQ
19 variation may arise from altered protein interactions, altered ELF3 nuclear localization,
20 altered regulation of the PIF4 developmental integrator, or a combination thereof.

21 Here, we dissect the epistatic landscape modifying the function of the *ELF3*-STR
22 through both physical and genetic interactions, and present evidence that this STR
23 forms the hub of a complex network of epistasis, likely due to its role as a compensatory
24 modifier of several other loci.

25 METHODS

26 **Plant material and growth conditions:** Hypocotyl length was assayed in seedlings
27 grown for 15d in incubators set to SD (8h light : 16h dark) at 22° on vertical plates as
28 described previously (Undurraga *et al.* 2012). The *elf3-200* (Undurraga *et al.* 2012) and
29 *elf3-4* (Hicks *et al.* 1996) mutants have been previously described. T-DNA lines (Alonso
30 *et al.* 2003; Kleinboelting *et al.* 2012) were obtained from the Arabidopsis Biological
31 Resource Center (Ohio State University).

32
33 **Genotyping:** For genotyping the *ELF3* STR and other loci across many F₂ segregants,
34 1-2 true leaves from each seedling were subjected to DNA extraction. Seedlings were
35 stored on their growth plates at 4° before genotyping but after phenotypic analysis. For
36 genotyping the *ELF3* STR, PCR was performed in 10 µL volume containing 0.5 µM
37 primers (Table S1), 0.2 µM each dNTP, 1 µL 10X ExTaq buffer, and 0.1 U ExTaq
38 (Takara, Tokyo, Japan); with initial denaturation step of 95° for 5', followed by 40 cycles
39 of 95° 30", 49° 20", 72° 10", with a final extension step at 72° for 5'. For other loci, PCR
40 was performed in 20 µL volume containing 0.5 µM primers (Table S1), 0.2 µM each
41 dNTP, 2 µL 10X ExTaq buffer, and 0.25 U Taq polymerase (NEB, Ipswich, MA); with
42 initial denaturation step of 95° for 5', followed by 35 cycles of 95° 30", 55° 30", 72° 1',
43 with a final extension step at 72° for 5'.
44

1 **Genome resequencing:** Plants selected for genotyping-by-sequencing were
2 transplanted to soil and grown under LD for 2-3 weeks. They were then stored at 4° until
3 DNA extraction was performed. One late rosette-stage Ws individual was used for Ws
4 whole-genome resequencing. DNA extraction was performed using the DNeasy Plant
5 Mini kit (Qiagen, Valencia, CA) according to the kit protocol. This DNA was quantified
6 using high-sensitivity Qubit fluorescence analysis (ThermoFisher Scientific, Waltham,
7 MA) and re-genotyped with *ELF3*-STR primers (Table S1). We used 10 ng DNA from
8 each F₂ segregant in NextEra transposase library preparations (Illumina, San Diego,
9 CA), or a standard 50 ng preparation for the Ws library. Library quality was assessed on
10 a BioAnalyzer (Agilent, Santa Clara, CA) or agarose gels. The Ws individual was
11 sequenced in one 300-cycle MiSeq v2 run (300 bp single-end reads) to ~12X coverage.
12 The F₂ segregant libraries were pooled and sequenced in one 200-cycle HiSeq v3 run
13 to ~2X average coverage (100 bp paired-end reads, Table S3).

14
15 **Sequence analysis:** Reads were aligned to the Col reference genome using BWA
16 v0.7.5 MEM (Li 2013), and variants were called using SAMtools v0.1.19 (Li *et al.* 2009).
17 High-quality Ws variants (Q_≥40) were thus identified from Ws parent data, and
18 compared with variants in previously-sequenced related strains (Gan *et al.* 2011). F₂
19 segregant genotype calls were combined into a single variant call format (VCF) file and
20 filtered for loci with such Ws variants. We used SNPtools (Wang *et al.* 2013) to perform
21 haplotype and genotype imputation for each locus in F₂ segregants. For workflows
22 employed in sequence analysis, see Supplementary Text. Following sequence analysis,
23 one individual was found to be a heterozygote at the *ELF3* locus. This individual was
24 omitted from all following analyses requiring *ELF3* homozygotes.

25
26 **Quantitative trait locus (QTL) analysis:** F₂ genotypes were reduced to a set of 500
27 loci randomly sampled from the imputed genotypes, plus a single nucleotide variant
28 (SNV) marking the *ELF3* locus. We used these genotypes to estimate a genetic map
29 and perform QTL analysis using the R/qtl package (Broman *et al.* 2003). A
30 nonparametric epistasis test was implemented in a custom R script using R/qtl
31 functions. For a more detailed description of commands and the epistasis test, see the
32 Supplementary Text. Follow-up genotyping of 10 additional F₂s was performed using
33 PCR markers (Table S1), and these genotypes were included in final QTL analyses.

34
35 **Candidate gene analysis:** Homozygous T-DNA lines (Alonso *et al.* 2003; Kleinboelting
36 *et al.* 2012) with insertions in genes of interest were obtained from the Arabidopsis
37 Biological Resource Center (Ohio State University) and phenotyped for hypocotyl length
38 under SD at 15 days. All such experiments were performed at least twice. Double
39 mutants were obtained by crossing relevant lines and genotyping (primers in Table S1).
40 Mutant lines are listed in Table S2. Expression analysis confirmed no detectable *LSH9*
41 expression in the *lsh9* mutant, suggesting that it is a null mutant. *LSH9* promoter
42 sequences across strains were downloaded from the Salk 1,001 Genomes Browser
43 (<http://signal.salk.edu/atg1001/3.0/gebrowser.php>) (Alonso-Blanco *et al.* 2016) and
44 aligned using Clustal Omega v1.0.3 (Sievers *et al.* 2011).

1
2 **Yeast two-hybrid (Y2H):** *ELF3* variants with different STR lengths were PCR cloned
3 out of cDNAs of previously described *A. thaliana* carrying *ELF3* transgenes (Undurraga
4 *et al.* 2012) into the *Xma*I/*Bam*HI sites of pGBKT7. Genes to be tested for *ELF3*
5 interactions were PCR cloned into the *Eco*RI/*Xho*I sites of pGADT7 from cDNAs of
6 indicated strains (Table S1 for primers). Clones were confirmed by restriction digest and
7 sequencing. The Y2H screen was performed against the Arabidopsis Mate and Plate
8 cDNA library (Clontech, Madison, WI), essentially according to the manufacturer's
9 instructions, except selections were performed on C-leu –trp –his plates incubated at
10 23°. Clones which also showed activation of the *ADE2* reporter gene and did not
11 autoactivate were subsequently tested against the various *ELF3*-polyQ constructs (see
12 Supplementary Text for details, full details on clones given in File S1).

13 LacZ activity was assayed through X-gal cleavage essentially as previously
14 described (Möckli and Auerbach 2004), again in strains using PJ69-4a as Mata parent.
15 For weakly activating constructs (GLDP1 and *ELF4*), 0.2 absorbance units of yeast
16 were used in each assay to reduce background, and color development was assessed
17 at points between 16 and 72 hours of incubation at room temperature.

18
19 **Quantitative PCR (qPCR):** For measuring *ELF3* and *LSH9* transcript levels, pooled
20 aerial tissue of ~30 mg short-day-grown seedlings of each relevant genotype were
21 collected at ZT8 7d post germination. RNA preparation and qPCR was performed as
22 described previously (Undurraga *et al.* 2012), using primers in Table S1.

23
24 **Statistical analysis:** All statistical analyses and plotting was performed using R 2.15.3
25 or R 3.2.1 (R Core Team 2016). Analysis scripts are provided with data as detailed
26 below.

27
28 **Data availability:** High-throughput sequencing data are available in BAM format at
29 NCBI Sequence Read Archive accession SRP077615. Processed genotype data,
30 phenotype data, and analysis code are available at
31 <https://figshare.com/s/e01a40b98a4ef5a9e5b3>.

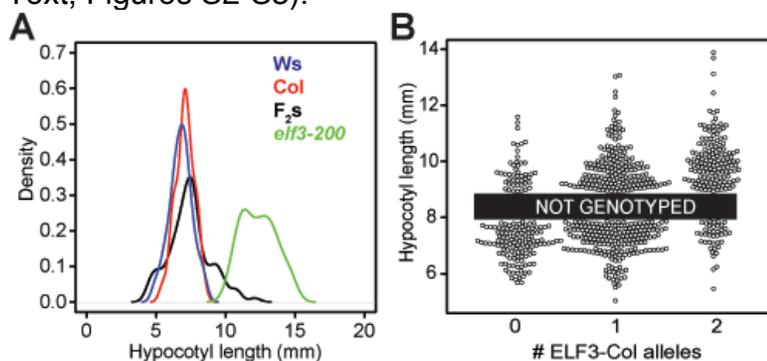
32 33 RESULTS

34 **Genetic analysis of *ELF3*-STR effects on hypocotyl length:** To investigate the
35 genetic architecture of epistasis for the *ELF3*-STR, we crossed two *A. thaliana* strains
36 with a previously reported mutual incompatibility of their respective *ELF3*-STR variants
37 (Col-0, *Ws*, (Undurraga *et al.* 2012)). We phenotyped the resulting F₂ population for
38 hypocotyl length under short days, a trait dramatically affected by *ELF3* function. The
39 Col and *Ws* backgrounds did not substantially differ in this trait ($p = 0.16$, Kolmogorov-
40 Smirnov test, Figure 1A). Although most F₂ seedlings showed phenotypes within the
41 range of the two parental lines, the F₂ phenotypic distribution showed a long upper tail
42 of transgressive variation, and consequently a different distribution from either parent (p
43 = 0.0039 against *Ws*, $p = 0.055$ against Col, Kolmogorov-Smirnov tests). As longer
44 hypocotyls in light conditions indicate *ELF3* dysfunction in the circadian clock (Liu *et al.*

1 2001), this observation is consistent with the co-segregation of incompatible Col and Ws
2 alleles. We replicated this observation in a much larger population (1106 seedlings),
3 which was used for further genetic analysis.

4
5 To investigate the genetic basis of the phenotypic transgression in hypocotyl length, we
6 harvested the 720 most phenotypically extreme seedlings (longest and shortest
7 hypocotyls) for genotyping (Figure S1). Each individual seedling was genotyped at the
8 *ELF3* locus, using primers directly ascertaining the 27bp *ELF3*-STR-length
9 polymorphism between Col and Ws. Across these individuals, we observed a strong
10 main effect of the *ELF3* locus on phenotype (Figure 1B), in which the Col allele of *ELF3*
11 frequently showed transgressive phenotypes, though some individuals homozygous for
12 the Ws allele also showed transgressive phenotypes. Specifically, a naïve regression
13 analysis of the data in Figure 1B indicated that each *ELF3*-Col allele increased
14 hypocotyl length by 0.87 ± 0.077 mm, and that the *ELF3* locus thereby explained 15% of
15 phenotypic variation. This analysis is misleading, because it implies that Col seedlings
16 should show longer hypocotyls than Ws seedlings due to *ELF3* genotype – this is not
17 the case (Figure 1A).

18 Among these seedlings, the individuals with extreme phenotypes and individuals
19 homozygous at the *ELF3* locus are expected to be most informative about *ELF3*-*STR*
20 effects on phenotype. Furthermore, *ELF3* genetic interactions are expected to be most
21 apparent in *ELF3* homozygotes. Consequently, we used a novel genetic approach to
22 detect epistasis between *ELF3* and other loci as follows. For each *ELF3* STR allele, we
23 selected 24 homozygotes (Ws/Ws and Col/Col) at each phenotypic extreme (the
24 shortest and longest hypocotyls). The sampling of extremes is an effective and
25 statistically justified method for genetic mapping (Lander and Botstein 1989). These 96
26 individuals were analyzed in a genotyping-by-sequencing approach (Table S3, Figure
27 S1). For details of this approach, see the Supplementary Information (Supplementary
28 Text, Figures S2-S3).

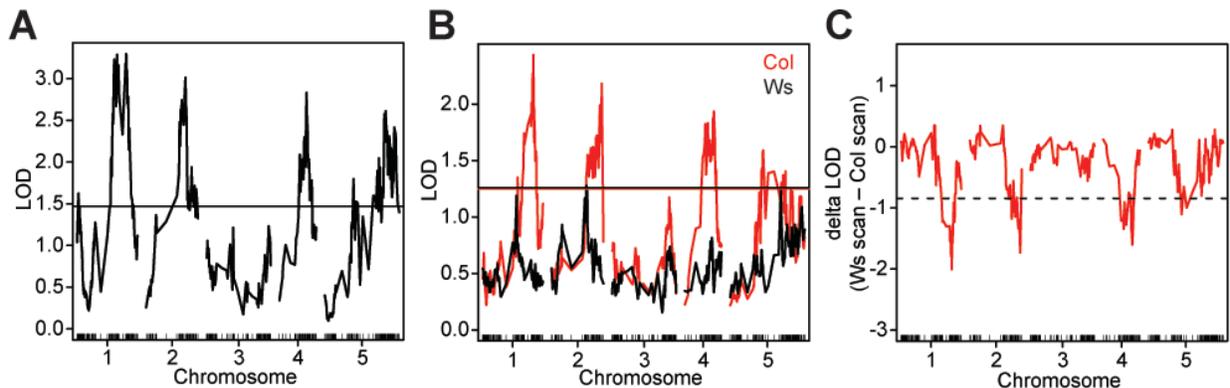


29
30 **Figure 1. Phenotypic transgression in Col x Ws F₂ segregants.** (A): F₂ segregant
31 phenotypes compared to parents and an *elf3* null mutant in the Col background; n ≈ 50
32 for each homozygous line and n ≈ 100 for F₂s. Colors indicate genotypes. Hypocotyl
33 length was determined at 15d under short days. (B): Phenotypic distributions of a large
34 population of Col x Ws F₂ segregants, for 720 extreme individuals genotyped at the
35 *ELF3* locus. N total = 1106 seedlings. 386 seedlings were not genotyped (indicated by
36 the black box).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

With these data, we performed a one-dimensional 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 potentially 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 binned F_2 s homozygous at *ELF3* according to their *ELF3* genotype, and performed one-dimensional QTL scans on each homozygote group separately (masking the genotypes of all other individuals). We observed that the same LOD peaks were replicated well in *ELF3*-Col homozygotes, but poorly in *ELF3*-Ws homozygotes (Figure 2B). Notably, a second Chr2 QTL was thus revealed, indicating that loci other than *ELF3* on Chr2 are relevant to the phenotype (at least in *ELF3*-Col plants). This analysis suggested that the *ELF3*-STR genotype is epistatic with at least four other loci controlling this phenotype, with effects masked in *ELF3*-Ws plants.



17
18
19
20
21
22
23
24
25
26

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. Dotted horizontal line indicates 99% significance threshold based on permutations.

27
28
29
30
31
32
33
34
35

To directly test for epistasis with *ELF3*, we adapted a previously described method (Sangster *et al.* 2008a). Separating the *ELF3* homozygotes again, we used permutations to define an empirical null distribution for the difference of likelihood (LOD) scores expected between the Ws and the Col scans. When comparing the difference in LOD scores between the two QTL scans at all loci, we found that the peaks on Chr1, Chr2, and Chr4 (and to a lesser extent Chr5) were all stronger in Col (Figure 2C). Consequently, these loci constitute background-specific *ELF3* interactors.

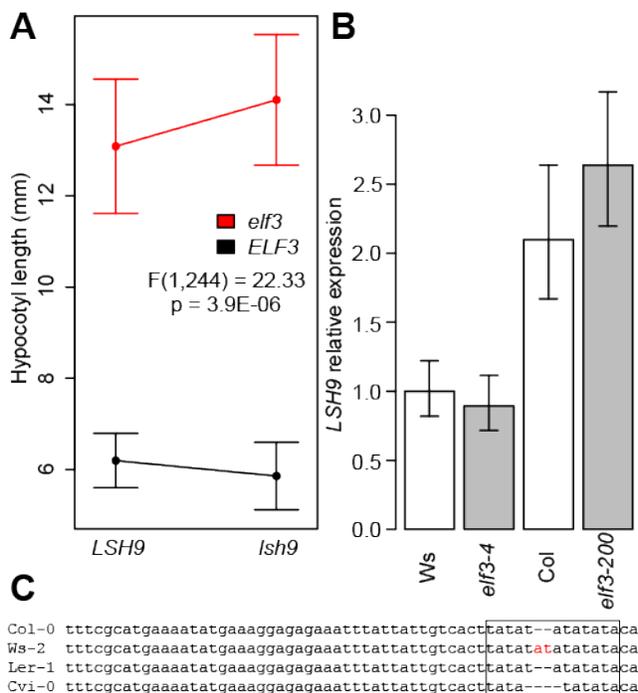
We considered the genetic contribution of these loci to the phenotype using a multiple QTL mapping approach, using both the independently estimated QTL locations

1 and a refined model re-estimating QTL positions based on information from all QTLs
 2 (Table S4). In each case, loci of strong effect on Chr1, Chr2, and Chr4 were supported,
 3 along with interactions between Chr2 (*ELF3*) and the other two loci. In the refined
 4 model, the Chr5 locus and the second (other than *ELF3*) Chr2 locus were also strongly
 5 supported. We conclude that although *ELF3* interacts epistatically with a variety of other
 6 loci in determining hypocotyl length, the principal contributors to *ELF3*-mediated effects
 7 on the trait are on Chr1 and Chr4. Moreover, direct inspection of phenotypic effects of
 8 *ELF3* in interaction with each putative locus among F₂ segregants supported the
 9 hypothesis of epistasis with *ELF3* most clearly for the Chr1 and Chr4 loci (Figure S5).

10

11 **Candidate gene analysis identifies *LSH9* as a genetic interactor of *ELF3*:** The
 12 chromosome intervals identified by our QTL analysis encompassed a large number of
 13 genes, and overlapped with an inversion between these backgrounds on Chr4 (Rowan
 14 *et al.* 2015). Previous work using a multiparent *A. thaliana* mapping population (Kover *et*
 15 *al.* 2009) also identified possible candidate genes affecting hypocotyl length in these
 16 regions (Khattak 2014).

17 We phenotyped mutants of several candidate genes in the Col background under
 18 the conditions of our intercross experiment (15d SD hypocotyl length, Figure S6). We
 19 observed small phenotypic effects of the T-DNA insertion mutants *Ish9* and *nup98*.
 20 However, these small effects on their own cannot explain the transgressive phenotypic
 21 variation in F₂S (Figure 1A).
 22



23

24 **Figure 3. *ELF3* interacts genetically with *LSH9*, which shows background-specific**
 25 **expression. (A):** Double mutant analysis of *elf3* and *lsh9* seedlings grown for 15d under
 26 SD. ANOVA analysis of the interaction between the mutant effects on phenotype is
 27 displayed. Error bars indicate standard deviation, n > 35 for each genotype. (B): qRT-

1 PCR analysis of *LSH9* transcript levels across genotypes. Seedlings were grown under
2 SD and harvested at ZT8 (dusk). *LSH9* expression is expressed as a proportion of *Ws*
3 expression, normalized relative to *UBC21*, error bars are standard error from three
4 technical replicates. Experiments were repeated with similar results. (C): *Ws*-specific
5 STR polymorphism in the *LSH9* promoter. The -221 to -163 region (relative to start
6 codon) of the *LSH9* promoter is aligned across diverse *A. thaliana* strains (Col, *Ws*-2,
7 Ler-1, Cvi-0), with the small [TA]_n STR boxed. *Ws*-2 is a separately maintained stock of
8 the *Ws* (Wassilewskija) strain. The *Ws*-2-specific polymorphism is highlighted in red.

9
10 We generated double mutants between these mutants and the *elf3* null mutant to
11 determine whether these genes interacted epistatically with *ELF3*. We found little
12 evidence for an interaction between *nup98* and *elf3* mutations (Figure S6C). However,
13 we detected a significant interaction between *ELF3* and *LSH9*, in the form of reciprocal
14 sign epistasis between the two null mutants affecting hypocotyl length (Figure 3A).
15 Although *lsh9* single mutants had significantly shorter hypocotyls than WT, *lsh9 elf3*
16 double mutant hypocotyls were substantially longer than in *elf3* single mutants. *LSH9*
17 (*LIGHT-DEPENDENT SHORT HYPOCOTYLS 9*) is an uncharacterized gene belonging
18 to a gene family named for LSH1, which is known to act in hypocotyl elongation (Zhao
19 *et al.* 2004). Like other genes in this family, *LSH9* encodes a putative nuclear
20 localization sequence but no other distinguishing features.

21 To test our hypothesis that *ELF3*-STR mediated epistasis may be due to altered
22 protein interactions, we investigated whether *LSH9* and *ELF3* interacted physically
23 using Y2H. However, we were unable to detect a physical interaction between the Col
24 or *Ws* variants of *LSH9* and *ELF3* (Figure S7), suggesting a different mechanistic basis
25 for the observed genetic interaction. For example, *LSH9* expression may depend on
26 *ELF3* function as a transcriptional regulator. Alternatively, *ELF3* expression may depend
27 on *LSH9* function. We tested both hypotheses by measuring expression levels of
28 *LSH9/ELF3* in, respectively, *elf3* or *lsh9* mutant backgrounds (*elf3* mutants were
29 available in both Col and *Ws* backgrounds, *lsh9* only in Col). *ELF3* expression levels
30 were unchanged in *lsh9* mutants (Figure S8). Moreover, levels of *LSH9* transcript did
31 not significantly differ between WT and *elf3* mutants in either strain background.
32 However, *LSH9* expression was reduced in the *Ws* background relative to Col
33 independently of *ELF3* genotype (Figure 3B). This result is consistent with the observed
34 phenotypic interaction in F₂s, which showed elongated hypocotyls when Col alleles at
35 the *ELF3* locus co-segregated with *Ws* alleles at the *LSH9* locus (Figure S5), thereby
36 pairing poorly-functioning *ELF3* alleles with potentially lower *LSH9* expression levels.

37 Taken together, *ELF3*-*LSH9* epistasis between Col and *Ws* may be due to
38 regulatory changes between these two backgrounds altering *LSH9* transcript levels.
39 Coincidentally, we observed that the *LSH9* promoter contains an STR polymorphism in
40 the *Ws* background that may alter *LSH9* expression; alternatively *LSH9* altered
41 expression in *Ws* may be due to trans-effects (Figure 3C).

42

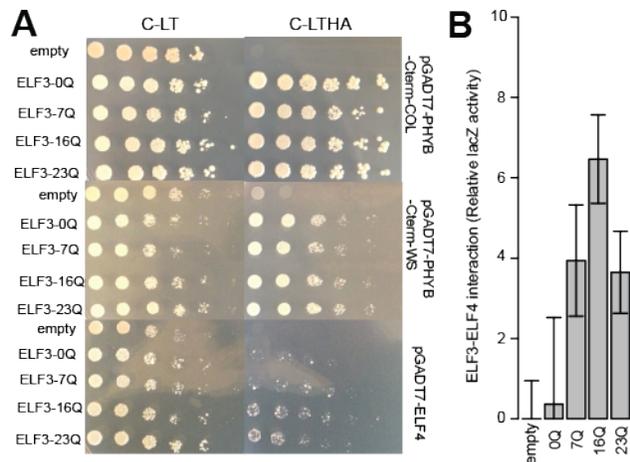


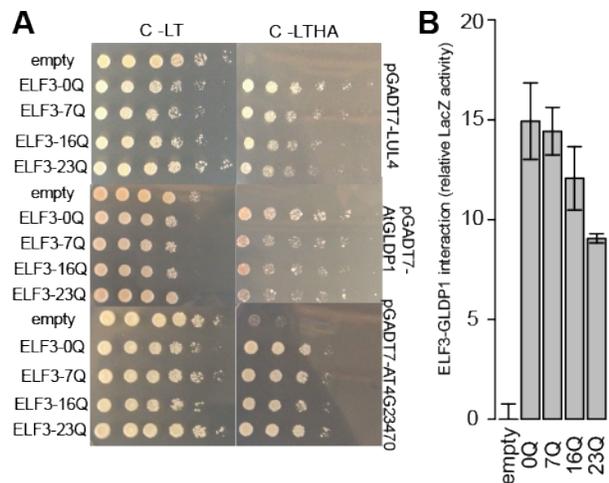
Figure 4. Y2H interaction of ELF3 with known protein interactors can be modulated by polyQ variation.

(A): 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. 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 cell lysates (see 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 standard deviation across three technical replicates. This experiment was repeated with similar results.

ELF3-polyQ tract variation affects known protein interactions: 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 synthetic ELF3s with 0Q (no polyQ), 7Q (variant in Col), 16Q (variant in Ws), and 23Q (endogenous to strains Br-0 and Bur-0) forms of ELF3 show Y2H interactions with well-described ELF3 interactors PHYB (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 auto-activation 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, though rather weakly compared to PHYB. Under these conditions, a subtle preference of ELF4 for longer polyQ variants (e.g. ELF3-16Q and ELF3-23Q) was apparent. We confirmed this preference in a quantitative, growth-independent assay in which LacZ expression is driven by the Y2H interaction (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.

1 Together, our data suggest that ELF3-polyQ tract variation can affect ELF3
2 protein interactions, in particular if these interactions are weaker (as for ELF4) and
3 presumably more sensitive to structural variation in ELF3.
4



5
6 **Figure 5. Y2H screen identifies new interactors of ELF3.** (A): Yeast carrying
7 indicated vectors were spotted in five-fold dilutions onto C-leu-trp (C-LT) or C-leu-trp-
8 his-ade (C-LTHA) media. For each protein X, experiments were repeated with at least
9 two independent PJ69-4α + pGADT7-X transformants with similar results. (B): LacZ
10 assays support polyQ effects on ELF3-At-GLDP1 interaction. The strains shown in (A)
11 also express *LacZ* from the Y2H promoter, whose activity was assayed in cell lysates
12 (see Methods). In each assay, all observations are expressed relative to the activity of
13 the empty vector, whose mean is set to 0. Error bars indicate standard deviation across
14 three technical replicates. This experiment was repeated with similar results.
15

16 **Y2H screen identifies three novel ELF3 interactors, one of which is polyQ-**
17 **modulated:** None of the known ELF3 interactors were encoded by genes located in the
18 major Chr1 and Chr4 QTLs identified by our genetic screen. If the ELF3-polyQ tract
19 mediates protein interactions, these regions should contain additional, previously-
20 undescribed polyQ-modulated ELF3 interactors. We screened the ELF3-7Q protein for
21 interactions with proteins from a commercially available library derived from Col, to
22 detect ELF3-protein interactions within the Col background.

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

1 all interactors. Indeed, the modest effects that we found were interactor-specific and
2 thus not likely to generalize.

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

10 Alternatively, rather than modulating ELF3 function as an encoded polyQ tract,
11 the *ELF3*-STR may affect ELF3 transcription or processing. A previous study of an
12 intronic STR in *A. thaliana* demonstrated that certain hyperexpanded STR alleles led to
13 dysregulation of the *ILL1* gene (Sureshkumar *et al.* 2009), presumably due to aberrant
14 processing of *ILL1* transcripts. Others have previously argued that such ‘informational’
15 (as opposed to ‘operational’) processes are more likely to have genetic or physical
16 interactions (Jain *et al.* 1999). We have not tested this hypothesis; however, our
17 previous studies found no correlation between *ELF3*-STR variation and *ELF3*
18 expression across many natural strains (Undurraga *et al.* 2012).

19 Taken together, our findings support a model in which highly variable STRs can
20 contribute to the epistatic component of heritability through both direct and indirect
21 functional interactions with other loci.

22

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

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

39 Next, the effects of the disease-associated polyQ hyperexpansions are generally
40 dominant, due to the nature of their molecular effects, which are generally thought to
41 show a protein gain-of-function (Orr and Zoghbi 2007; Fondon *et al.* 2008). We
42 observed no evidence that ELF3-polyQ variation behaves in a dominant fashion, but
43 rather that Col x Ws F₁s show approximately WT phenotypes. The effects of ELF3-

1 polyQ variation manifest only when separated from a favorably interacting genetic
2 context (as in segregating F₂s).

3 Last, the (sometimes implicit) expectation from polyQ hyperexpansion
4 phenotypes is that there is a linear, or at least monotonic, association between
5 phenotypes and polyQ length. For instance, the degree of huntingtin polyQ expansion is
6 strongly correlated with Huntington's disease severity (ANDREW *et al.* 1993). In
7 contrast, no ELF3-polyQ-related phenotype has been shown to have a monotonic
8 association to ELF3 polyQ length (Undurraga *et al.* 2012; Press *et al.* 2014, 2016).
9 Indeed, all indications are that the mapping between ELF3 polyQ tract length and
10 phenotype is non-monotonic and strongly contingent on genetic background, unlike the
11 classic polyQ disease models. Our results suggest that sub-expansion polyQ tract
12 variation can engage in multiple genetic interactions, and at least in some cases
13 modulate protein interactions. More work is needed to evaluate the generality of our
14 findings and determine the breadth of molecular mechanisms by which modest, sub-
15 expansion polyQ variation can affect phenotype.

16
17 **ELF3 as a robustness gene:** Here, we operated under the assumption that a few
18 strong polyQ-modulated interactions with ELF3 explain the polyQ-dependent genetic
19 architectures. Alternatively, the ELF3 polyQ tract may modulate many transient
20 interactions that are perturbed by hypomorphic ELF3 activity. In this interpretation, *ELF3*
21 acts as a 'robustness gene' (Lempe *et al.* 2013). The best-described example of such is
22 the protein chaperone HSP90 (Rutherford and Lindquist 1998), whose multiple transient
23 interactions with many proteins—about 10% of the yeast proteome (Zhao *et al.* 2005)—
24 lead to pleiotropic effects upon HSP90 inhibition or dysregulation (Sangster *et al.* 2007).
25 ELF3 has been previously proposed as a robustness gene (Jimenez-Gomez *et al.*
26 2011), consistent with its promiscuity in protein complexes and the pleiotropic nature of
27 *elf3* mutant phenotypes. Our finding that functional modulation of ELF3 by polyQ
28 variation reveals several genetic interactors is consistent with this interpretation.

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

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

41

42 **ACKNOWLEDGMENTS**

43 We thank Karla Schultz and Katie Uckele for technical assistance. We thank Choli Lee
44 and Jay Shendure for assistance with high-throughput sequencing of the Col x Ws F₂

1 population and members of the Shendure laboratory for advice regarding library
2 preparation. We thank Amy Lanctot for generating the pGBK-ELF3-0Q and pGBK-
3 ELF3-23Q constructs. We thank Daniel Melamed and Stanley Fields for guidance in
4 carrying out Y2H experiments and the generous gift of yeast strains. We thank Giang
5 Ong and Maitreya Dunham for access to the MiSeq instrument for resequencing the Ws
6 genome. We thank Stanley Fields and Evan Eichler for access to LightCycler
7 instruments. We thank Elhanan Borenstein and members of the Queitsch and
8 Borenstein laboratories for helpful conversations. MOP was supported in part by
9 National Human Genome Research Institute Interdisciplinary Training in Genome
10 Sciences Grant 2T32HG35-16. CQ is supported by National Institute of Health New
11 Innovator Award DP2OD008371.

12

13

REFERENCES

14

15 Alon U., 2003 Biological networks: the tinkerer as an engineer. *Science* **301**: 1866–7.

16 Alonso J. M., Stepanova A. N., Leisse T. J., Kim C. J., Chen H., Shinn P., Stevenson D. K.,

17 Zimmerman J., Barajas P., Cheuk R., Gadrinab C., Heller C., Jeske A., Koesema E.,

18 Meyers C. C., Parker H., Prednis L., Ansari Y., Choy N., Deen H., Geralt M., Hazari N.,

19 Hom E., Karnes M., Mulholland C., Ndubaku R., Schmidt I., Guzman P., Aguilar-

20 Henonin L., Schmid M., Weigel D., Carter D. E., Marchand T., Risseeuw E., Brogden D.,

21 Zeko A., Crosby W. L., Berry C. C., Ecker J. R., 2003 Genome-wide insertional

22 mutagenesis of *Arabidopsis thaliana*. *Science* **301**: 653–7.

23 Alonso-Blanco C., Andrade J., Becker C., Bemm F., Bergelson J., Borgwardt K. M., Cao J., Chae

24 E., Dezwaan T. M., Ding W., Ecker J. R., Exposito-Alonso M., Farlow A., Fitz J., Gan X.,

25 Grimm D. G., Hancock A. M., Henz S. R., Holm S., Horton M., Jarsulic M., Kerstetter R.

26 A., Korte A., Korte P., Lanz C., Lee C.-R., Meng D., Michael T. P., Mott R., Mulyati N. W.,

27 Nägele T., Nagler M., Nizhynska V., Nordborg M., Novikova P. Y., Picó F. X., Platzer A.,

28 Rabanal F. A., Rodriguez A., Rowan B. A., Salomé P. A., Schmid K. J., Schmitz R. J.,

29 Seren Ü., Sperone F. G., Sudkamp M., Svoldal H., Tanzer M. M., Todd D., Volchenboum

- 1 S. L., Wang C., Wang G., Wang X., Weckwerth W., Weigel D., Zhou X., 2016 1,135
2 Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis thaliana*. *Cell* **0**.
- 3 Anwer M. U., Boikoglou E., Herrero E., Hallstein M., Davis A. M., James G. V., Nagy F., Davis S.
4 J., 2014 Natural variation reveals that intracellular distribution of ELF3 protein is
5 associated with function in the circadian clock. *eLife* **3**: e02206.
- 6 Broman K. W., Wu H., Sen S., Churchill G. A., 2003 R/qtl: QTL mapping in experimental
7 crosses. *Bioinformatics* **19**: 889–890.
- 8 Chow B. Y., Helfer A., Nusinow D. A., Kay S. A., 2012 ELF3 recruitment to the PRR9 promoter
9 requires other Evening Complex members in the *Arabidopsis* circadian clock. *Plant*
10 *Signal. Behav.* **7**: 170–3.
- 11 Escher D., Bodmer-Glavas M., Barberis A., Schaffner W., 2000 Conservation of Glutamine-
12 Rich Transactivation Function between Yeast and Humans. *Mol. Cell. Biol.* **20**: 2774–
13 2782.
- 14 Fondon J. W., Hammock E. A. D., Hannan A. J., King D. G., 2008 Simple sequence repeats:
15 genetic modulators of brain function and behavior. *Trends Neurosci.* **31**: 328–34.
- 16 Gan X., Stegle O., Behr J., Steffen J. G., Drewe P., Hildebrand K. L., Lyngsoe R., Schultheiss S. J.,
17 Osborne E. J., Sreedharan V. T., Kahles A., Bohnert R., Jean G., Derwent P., Kersey P.,
18 Belfield E. J., Harberd N. P., Kemen E., Toomajian C., Kover P. X., Clark R. M., Rättsch
19 G., Mott R., 2011 Multiple reference genomes and transcriptomes for *Arabidopsis*
20 *thaliana*. *Nature* **477**: 419–23.

- 1 Gemayel R., Vincés M. D., Legendre M., Verstrepen K. J., 2010 Variable tandem repeats
2 accelerate evolution of coding and regulatory sequences. *Annu. Rev. Genet.* **44**: 445–
3 77.
- 4 Hannan A. J., 2010 Tandem repeat polymorphisms: modulators of disease susceptibility
5 and candidates for “missing heritability”. *Trends Genet.* **26**: 59–65.
- 6 Herrero E., Kolmos E., Bujdosó N., Yuan Y., Wang M., Berns M. C., Uhlworm H., Coupland G.,
7 Saini R., Jaskolski M., Webb A., Gonçalves J., Davis S. J., 2012 EARLY FLOWERING4
8 recruitment of EARLY FLOWERING3 in the nucleus sustains the Arabidopsis
9 circadian clock. *Plant Cell* **24**: 428–43.
- 10 Hicks K. A., Millar A. J., Carre I. A., Somers D. E., Straume M., Meeks-Wagner D. R., Kay S. A.,
11 1996 Conditional Circadian Dysfunction of the Arabidopsis early-flowering 3
12 Mutant. *Science* **274**: 790–792.
- 13 Jacob F., 1977 Evolution and Tinkering. *Science* **196**: 1161–1166.
- 14 Jain R., Rivera M. C., Lake J. a, 1999 Horizontal gene transfer among genomes: the
15 complexity hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* **96**: 3801–6.
- 16 Jimenez-Gomez J. M., Corwin J. a., Joseph B., Maloof J. N., Kliebenstein D. J., 2011 Genomic
17 Analysis of QTLs and Genes Altering Natural Variation in Stochastic Noise (G Gibson,
18 Ed.). *PLoS Genet.* **7**: e1002295.
- 19 Kashi Y., King D., Soller M., 1997 Simple sequence repeats as a source of quantitative
20 genetic variation. *Trends Genet.* **13**: 74–78.

- 1 Khattak A. K., 2014 Natural Variation in *Arabidopsis thaliana* Growth in Response to
2 Ambient Temperatures: PhD Thesis.
- 3 Kleinboelting N., Huet G., Kloetgen A., Viehoveer P., Weisshaar B., 2012 GABI-Kat
4 SimpleSearch: new features of the *Arabidopsis thaliana* T-DNA mutant database.
5 *Nucleic Acids Res.* **40**: D1211–5.
- 6 Kolmos E., Herrero E., Bujdosó N., Millar A. J., Tóth R., Gyula P., Nagy F., Davis S. J., 2011 A
7 Reduced-Function Allele Reveals That EARLY FLOWERING3 Repressive Action on
8 the Circadian Clock Is Modulated by Phytochrome Signals in *Arabidopsis*. *Plant Cell*
9 Online **23**: 3230–3246.
- 10 Kover P. X., Valdar W., Trakalo J., Scarcelli N., Ehrenreich I. M., Purugganan M. D., Durrant C.,
11 Mott R., 2009 A Multiparent Advanced Generation Inter-Cross to Fine-Map
12 Quantitative Traits in *Arabidopsis thaliana*. *PLoS Genet* **5**:
13 e1000551.
- 14 Lachowiec J., Queitsch C., Kliebenstein D. J., 2015 Molecular mechanisms governing
15 differential robustness of development and environmental responses in plants. *Ann.*
16 *Bot.*: mcv151.
- 17 Lander E. S., Botstein D., 1989 Mapping Mendelian Factors Underlying Quantitative Traits
18 Using RFLP Linkage Maps. *Genetics* **121**: 185–199.
- 19 Lempe J., Lachowiec J., Sullivan A. M., Queitsch C., 2013 Molecular mechanisms of
20 robustness in plants. *Curr. Opin. Plant Biol.* **16**: 62–9.

- 1 Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin
2 R., 2009 The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.*
3 **25**: 2078–9.
- 4 Li H., 2013 Aligning sequence reads, clone sequences and assembly contigs with BWA-
5 MEM. *ArXiv*: 3.
- 6 Liu X. L., Covington M. F., Fankhauser C., Chory J., Wagner D. R., 2001 ELF3 encodes a
7 circadian clock-regulated nuclear protein that functions in an Arabidopsis PHYB
8 signal transduction pathway. *Plant Cell* **13**: 1293–304.
- 9 Lu X., Murphy R. M., 2015 Asparagine Repeat Peptides: Aggregation Kinetics and
10 Comparison with Glutamine Repeats. *Biochemistry (Mosc.)* **54**: 4784–94.
- 11 Möckli N., Auerbach D., 2004 Quantitative β -galactosidase assay suitable for high-
12 throughput applications in the yeast two-hybrid system. *BioTechniques* **36**: 872–
13 876.
- 14 Nieto C., López-Salmerón V., Davière J.-M., Prat S., 2014 ELF3-PIF4 Interaction Regulates
15 Plant Growth Independently of the Evening Complex. *Curr. Biol.* **25**: 187–193.
- 16 Nusinow D. A., Helfer A., Hamilton E. E., King J. J., Imaizumi T., Schultz T. F., Farré E. M., Kay
17 S. A., 2011 The ELF4-ELF3-LUX complex links the circadian clock to diurnal control
18 of hypocotyl growth. *Nature* **475**: 398–402.
- 19 Orr H. T., Zoghbi H. Y., 2007 Trinucleotide Repeat Disorders. *Annu. Rev. Neurosci.* **30**: 575–
20 621.

- 1 Perutz M. F., Johnson T., Suzuki M., Finch J. T., 1994 Glutamine repeats as polar zippers:
2 their possible role in inherited neurodegenerative diseases. *Proc. Natl. Acad. Sci.* **91**:
3 5355–5358.
- 4 Press M. O., Carlson K. D., Queitsch C., 2014 The overdue promise of short tandem repeat
5 variation for heritability. *Trends Genet.* **30**: 504–512.
- 6 Press M. O., Lanctot A., Queitsch C., 2016 ELF3 polyQ variation in *Arabidopsis thaliana*
7 reveals a PIF4-independent role in thermoresponsive flowering. *bioRxiv*: 038257.
- 8 Queitsch C., Carlson K. D., Girirajan S., 2012 Lessons from model organisms: phenotypic
9 robustness and missing heritability in complex disease. (SM Rosenberg, Ed.). *PLoS*
10 *Genet.* **8**: e1003041.
- 11 R Core Team, 2016 *R: A Language and Environment for Statistical Computing*. R Foundation
12 for Statistical Computing, Vienna, Austria.
- 13 Rowan B. A., Patel V., Weigel D., Schneeberger K., 2015 Rapid and inexpensive whole-
14 genome genotyping-by-sequencing for crossover localization and fine-scale genetic
15 mapping. *G3 Bethesda Md* **5**: 385–98.
- 16 Rutherford S. L., Lindquist S., 1998 Hsp90 as a capacitor for morphological evolution.
17 *Nature* **396**: 336–42.
- 18 Sangster T. a, Bahrami A., Wilczek A., Watanabe E., Schellenberg K., McLellan C., Kelley A.,
19 Kong S. W., Queitsch C., Lindquist S., 2007 Phenotypic diversity and altered

- 1 environmental plasticity in *Arabidopsis thaliana* with reduced Hsp90 levels. *PloS*
2 *One* **2**: e648.
- 3 Sangster T. a, Salathia N., Lee H. N., Watanabe E., Schellenberg K., Morneau K., Wang H.,
4 Undurruga S., Queitsch C., Lindquist S., 2008a HSP90-buffered genetic variation is
5 common in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 2969–74.
- 6 Sangster T. a, Salathia N., Undurruga S., Milo R., Schellenberg K., Lindquist S., Queitsch C.,
7 2008b HSP90 affects the expression of genetic variation and developmental stability
8 in quantitative traits. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 2963–8.
- 9 Schaefer M. H., Wanker E. E., Andrade-Navarro M. A., 2012 Evolution and function of
10 CAG/polyglutamine repeats in protein-protein interaction networks. *Nucleic Acids*
11 *Res.* **40**: 4273–87.
- 12 Sharma D., Sharma S., Pasha S., Brahmachari S. K., 1999 Peptide models for inherited
13 neurodegenerative disorders: conformation and aggregation properties of long
14 polyglutamine peptides with and without interruptions. *FEBS Lett.* **456**: 181–185.
- 15 Sievers F., Wilm A., Dineen D., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H.,
16 Remmert M., Söding J., Thompson J. D., Higgins D. G., 2011 Fast, scalable generation
17 of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst.*
18 *Biol.* **7**: 539.

- 1 Stott K., Blackburn J. M., Butler P. J., Perutz M., 1995 Incorporation of glutamine repeats
2 makes protein oligomerize: implications for neurodegenerative diseases. Proc. Natl.
3 Acad. Sci. U. S. A. **92**: 6509–13.
- 4 Sureshkumar S., Todesco M., Schneeberger K., Harilal R., Balasubramanian S., Weigel D.,
5 2009 A genetic defect caused by a triplet repeat expansion in *Arabidopsis thaliana*.
6 Science **323**: 1060–3.
- 7 Szamecz B., Boross G., Kalapis D., Kovács K., Fekete G., Farkas Z., Lázár V., Hrtyan M.,
8 Kemmeren P., Groot Koerkamp M. J. A., Rutkai E., Holstege F. C. P., Papp B., Pál C.,
9 2014 The Genomic Landscape of Compensatory Evolution. PLoS Biol. **12**: e1001935.
- 10 Taipale M., Jarosz D. F., Lindquist S., 2010 HSP90 at the hub of protein homeostasis:
11 emerging mechanistic insights. Nat. Rev. Mol. Cell Biol. **11**: 515–28.
- 12 Tajima T., Oda A., Nakagawa M., Kamada H., Mizoguchi T., 2007 Natural variation of
13 polyglutamine repeats of a circadian clock gene *ELF3* in *Arabidopsis*. Plant
14 Biotechnol. **24**: 237–240.
- 15 Undurraga S. F., Press M. O., Legendre M., Bujdoso N., Bale J., Wang H., Davis S. J., Verstrepen
16 K. J., Queitsch C., 2012 Background-dependent effects of polyglutamine variation in
17 the *Arabidopsis thaliana* gene *ELF3*. Proc. Natl. Acad. Sci. U. S. A. **109**: 19363–19367.
- 18 Usdin K., 2008 The biological effects of simple tandem repeats: Lessons from the repeat
19 expansion diseases. Genome Res. **18**: 1011–1019.

- 1 Wang Y., Lu J., Yu J., Gibbs R. A., Yu F., 2013 An integrative variant analysis pipeline for
2 accurate genotype/haplotype inference in population NGS data. *Genome Res.* **23**:
3 833–42.
- 4 Yoshida R., Fekih R., Fujiwara S., Oda A., Miyata K., Tomozoe Y., Nakagawa M., Niinuma K.,
5 Hayashi K., Ezura H., Coupland G., Mizoguchi T., 2009 Possible role of early flowering
6 3 (ELF3) in clock-dependent floral regulation by short vegetative phase (SVP) in
7 *Arabidopsis thaliana*. *New Phytol.* **182**: 838–50.
- 8 Yu J.-W., Rubio V., Lee N.-Y., Bai S., Lee S.-Y., Kim S.-S., Liu L., Zhang Y., Irigoyen M. L.,
9 Sullivan J. A., Zhang Y., Lee I., Xie Q., Paek N.-C., Deng X. W., 2008 COP1 and ELF3
10 control circadian function and photoperiodic flowering by regulating GI stability.
11 *Mol. Cell* **32**: 617–30.
- 12 Zhao L., Nakazawa M., Takase T., Manabe K., Kobayashi M., Seki M., Shinozaki K., Matsui M.,
13 2004 Overexpression of LSH1, a member of an uncharacterised gene family, causes
14 enhanced light regulation of seedling development. *Plant J.* **37**: 694–706.
- 15 Zhao R., Davey M., Hsu Y.-C., Kaplanek P., Tong A., Parsons A. B., Krogan N., Cagney G., Mai
16 D., Greenblatt J., Boone C., Emili A., Houry W. A., 2005 Navigating the chaperone
17 network: an integrative map of physical and genetic interactions mediated by the
18 hsp90 chaperone. *Cell* **120**: 715–27.

19