Background-dependent effects of polyglutamine variation in the Arabidopsis thaliana gene ELF3

Soledad Francisca Undurraga*, Maximilian Oliver Press*, Matthieu Legendre*, Nora Bujdoso*, Jacob Bale*#,
Hui Wang†, Seth J. Davis†, Kevin J. Verstrepen*,# and Christine Queitsch*,#

*Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA 98195; †Centre National de la Recherche Scientifique, IGS Unité Mixte de Recherche 7256, Structural and Genomic Information Laboratory, Mediterranean Institute of Microbiology, Aix-Marseille University, FR-13288 Marseille, France; ‡Department of Plant Developmental Biology, Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany; §Molecular and Cellular Biology Program and †Department of Biochemistry, University of Washington, Seattle, WA 98195; ¶FAS Center for Systems Biology, Harvard University, Boston, MA 02138; Laboratory for Systems Biology, Vlaams Instituut voor Biotechnologie, Bio-Incubator, B-3001 Leuven, Belgium; and ¶Laboratory for Genetics and Genomics, Centre of Microbial and Plant Genetics, Department of Microbial and Molecular Systems, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium

Edited by Detlef Weigel, Max Planck Institute for Developmental Biology, Tübingen, Germany, and approved October 9, 2012 (received for review July 2, 2012)

Tandem repeats (TRs) have extremely high mutation rates and are often considered to be neutrally evolving DNA. However, in coding regions, TR copy number mutations can significantly affect phenotype and may facilitate rapid adaptation to new environments. In several human genes, TR copy number mutations that expand polyglutamine (polyQ) tracts beyond a certain threshold cause incurable neurodegenerative diseases. PolyQ-containing proteins exist at a considerable frequency in eukaryotes, yet the phenotypic consequences of natural variation in polyQ tracts that are not associated with disease remain largely unknown. Here, we use Arabidopsis thaliana to dissect the phenotypic consequences of natural variation in the polyQ tract encoded by EARLY FLOWERING 3 (ELF3), a key developmental gene. Changing ELF3 polyQ tract length affected complex ELF3-dependent phenotypes in a striking and nonlinear manner. Some natural ELF3 polyQ variants phenocopied elf3 loss-of-function mutants in a common reference background, although they are functional in their native genetic backgrounds. To test the existence of background-specific modifiers, we compared the phenotypic effects of ELF3 polyQ variants between two divergent backgrounds, Col and Ws, and found dramatic differences. In fact, the Col-ELF3 allele, encoding the shortest known ELF3 polyQ tract, was haploinsufficient in Ws × Col F1 hybrids. Our data support a model in which variable polyQ tracts drive adaptation to internal genetic environments.

In coding regions, tandem repeat (TR) copy number variation can have profound phenotypic effects (1). For example, TR copy number mutations that expand polyglutamine (polyQ) tracts past a threshold number of glutamines can cause incurable neurodegenerative diseases, such as Huntington disease and Spinocerebellar Ataxias (2, 3). PolyQ tract length correlates with onset and severity of polyQ expansion disorders, but for intermediate polyQ tracts this correlation is far weaker (4–8), suggesting that genetic and environmental modifiers exist (9–12). Despite their potential for pathogenicity, variable polyQ tracts occur frequently in eukaryotic proteins, many of them functioning in development and transcription (1, 13–15). Model organism studies have suggested that coding TRs are an important source of quantitative genetic variation that facilitates evolutionary adaptation (1, 16–19). For example, TR copy number variation in the yeast gene FLO1 correlates linearly with flocculation (20), a phenotype that is important for stress survival (17). As polyQ tracts often mediate protein interactions (2, 3, 21), polyQ-encoding TR copy number mutations could produce large and possibly adaptive phenotypic shifts.

To determine the phenotypic impact of naturally occurring polyQ variation (18, 22, 23) in a genetically tractable model, we focused on the gene EARLY FLOWERING 3 (ELF3), which encodes a polyQ tract that is highly variable across divergent Arabidopsis thaliana strains (accessions) (19, 24). ELF3 is a core component of the circadian clock and a potent repressor of flowering, and is considered a “hub protein” for its many interactions with various proteins (24–31). Consequently, elf3 loss-of-function mutants show pleiotropic phenotypes: they flower early, show poor circadian function, and grow long embryonic stems (hypocotyls) in light (25–27, 29, 30, 32). SNPs in ELF3 affect shade avoidance, a fitness-relevant plant trait (24, 33). ELF3 polyQ variation has been suggested to correlate with two parameters of the circadian clock: period and phase (19). The ELF3 polyQ tract may mediate ELF3 membership in protein complexes, although thus far no ELF3-binding protein is known to bind it (26, 28–30). We discovered that altering polyQ tract length has dramatic effects on ELF3-dependent phenotypes and that these effects are dependent on genetic background.

Results

ELF3-TR Variation Affects ELF3-Dependent Phenotypes. Among 181 natural A. thaliana accessions, the ELF3-TR encoded between 7 and 29Q (Fig. S1A and Table S1). For comparison, polyQ expansions over 20Q are associated with disease in the context of the SCAG gene, although most other disease-associated polyQ expansions are longer (2, 19, 24). The most frequent ELF3-TR encoded 16Q, whereas the shortest TR (7Q) was found in the reference strain Col-0. We set out to test whether naturally occurring ELF3-TR alleles affect ELF3-dependent phenotypes, and whether they do so in a linear manner as suggested by association studies (19) and found for coding TR variation in other genes (16, 20). We generated expression-matched transgenic lines for most natural ELF3-TR alleles in the loss-of-function elf3-4 mutant (Ws background, Fig. S1C and Tables S2 and S3) (32) and measured their flowering time and circadian clock-related phenotypes (Fig. 1 and Fig. S2 A–G). ELF3-TR variation significantly affected ELF3-dependent phenotypes, but there was no evidence of a linear relationship. The different ELF3-TR alleles resulted in phenotypes ranging from nearly full complementation of elf3-4 to nearly phenocopying the loss-of-function mutant. We used principal components analysis (PCA) to describe the complex effects of ELF3-TR alleles on all tested ELF3-dependent phenotypes


The authors declare no conflict of interest.

1 Deceased October 7, 2006.

*To whom correspondence should be addressed. E-mail: queitsch@u.washington.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1211021109/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1211021109
plants with functional ELF3 to show short hypocotyls, late flowering, increased rosette leaf number, and short petioles (Fig. 1 B–D and Fig. S2). The endogenous ELF3-16Q allele complemented both the early-flowering and long-hypocotyl phenotypes of elf3-4 (Fig. 1 B–D and Fig. S2). In contrast, both the long ELF3-23Q and the short ELF3-7Q allele (endogenous TR alleles in Br-0/Bur-0 and Col-0, respectively) behaved similarly to the elf3-4 loss-of-function allele (Fig. 1 B–D and Fig. S2), although they are functional in their native backgrounds. Neither Col-0 nor Br-0 and Bur-0 show the phenotypic characteristics of elf3-4 mutants (early flowering (34), long hypocotyls (35), and long petioles (36)), suggesting that ELF3-TR alleles may interact with background-specific modifiers. ELF3-0Q, an artificial ELF3 allele lacking the TR, partially complemented elf3-4 (Fig. L4 and Fig. S2). Hence, the polyQ-encoding TR is not necessary for all ELF3 function, but changes in TR copy number are sufficient to enhance or ablate ELF3 function.

PC2 separated ELF3-20Q and ELF3-11Q, which behaved as hypomorphs in certain phenotypes but not others (Fig. L4). For example, ELF3-20Q plants had significantly longer hypocotyls than wild-type and its petioles phenocopied the extremely long petioles of the elf3-4 mutant (Fig. 1 C–E), but they did not differ from wild-type in flowering time (days to flower, Fig. 1B). The existence of both general and specific hypomorphs suggests that polyQ variation affects the multiple ELF3 functions separately.

As part of a protein complex, ELF3 affects expression of Phytochrome-interacting factor 5 (PIF5) and Pseudoresponsive regulator 9 (PRR9) (28, 37, 38). ELF3-16Q phenocopied wild-type PRR9 and PIF5 expression, and the hypomorphic ELF3-23Q phenocopied elf3-4 (28, 37, 38), mirroring their developmental phenotypes. Consistent with their divergence along PC2 (Fig. 1A), ELF3-11Q and ELF3-20Q differed in their effect on PRR9 expression, but not on PIF5 expression (Fig. S3 A and B), demonstrating that ELF3 polyQ variation differentially affects the regulation of downstream genes.

**ELF3-TR Variation Modulates the Precision of the Circadian Clock.** To directly assess the role of ELF3 polyQ variation in the circadian clock, we used the CCR2::LUC reporter system (25, 39). We observed little difference in circadian period among wild-typeWs and tested ELF3-TR alleles (Fig. S4–A). This contradicts a previously observed association of TR copy number with period in natural accessions (19). However, we found that the relative amplitude error (RAE) of oscillation varies substantially across ELF3-TR genotypes (Fig. 2A and Fig. S4B). RAE measures the precision of a circadian period (40): high RAE values (>0.4) indicate poor oscillation and clock dysfunction (41). The endogenous Ws ELF3-16Q nearly complemented the elf3-4 RAE defect, whereas the TR alleles ELF3-7Q, ELF3-10Q, and ELF3-23Q showed higher RAE, approaching arrhythmic elf3-4 levels (Fig. 2A and B), consistent with their hypomorphic performance in other ELF3 traits (close to elf3-4 in PC1 (Fig. L4)). Taken together, these results suggest that ELF3 polyQ tract length is a critical determinant of circadian clock precision—but not period length—in A. thaliana.

**ELF3-TR Variation Interacts with Genetic Background.** To test our hypothesis that ELF3-TR variation interacts with genetic background, we regenerated all ELF3-TR transgenic lines in the elf3-200 loss-of-function mutant with matched transgene expression (Col background; Fig. S1D and Table S4) (42). We used PCA to compare ELF3-TR effects between Ws and Col backgrounds (Fig. 3A and Fig. S5).

The Col-specific ELF3-7Q allele complemented elf3-200 in some traits, such as flowering time (in short days, SD) and hypocotyl length (in long days, LD), but not others (Fig. 3 A and B, and Figs. S5 and S6). This result may be because of the absence...
of the small 5’ intron from the ELF3 construct used in this study. However, there was still a dramatic spread of phenotypes: all longer ELF3-TR alleles (>20 Qs) nearly complemented elf3-200, delaying flowering and shortening hypocotyls, whereas few of the shorter alleles did (Fig. 3, and Figs. S5 and S6). Results were similar when the Col data were analyzed alone (Fig. S6). Thus, in contrast to our results in the Ws background, ELF3-TRs appeared to show a threshold effect for TR copy number in the Col background. We speculate that the intensive laboratory propagation of the Col-0 accession may have altered selection on the ELF3-TR, resulting in an extremely short “hypomorphic” allele, whereas under natural conditions a longer TR might be more functional.

Comparing TR allele effects between the two backgrounds revealed striking differences. For example, the ELF3-23Q allele was a general hypomorph in the Ws background (elf3-4), whereas it produced highly functional ELF3 in the Col background (elf3-200) (Fig. 3). In turn, the ELF3-16Q allele produced highly functional ELF3 in the Ws background (elf3-4), but was a general hypomorph in the Col background (elf3-200). The consistent performance of the artificial ELF3-0Q allele across backgrounds suggests that the background effect is TR-dependent (Fig. 3A and Fig. S5). Collectively, our results suggest that ELF3-TR alleles interact with background-specific modifiers.

**Col ELF3 Allele Is Haploinsufficient in Col x Ws Hybrids.** To address whether Ws and Col-specific background effects are sufficient for altered hybrid phenotypes, we generated F1 populations between wild-type and elf3-null plants in the Ws and Col backgrounds and measured ELF3 function by assessing hypocotyl length. Ws × Col F1 hybrids resembled their wild-type parents (Fig. 4). F1 hybrids containing both loss-of-function alleles had significantly longer hypocotyls than either parent (Fig. 4). Both ELF3 alleles were haploinsufficient in F1 crosses within their native backgrounds, as expected for recessive mutants (Fig. 4). In stark contrast, we observe that ELF3-Col, but not ELF3-Ws, phenocopied the extreme hypocotyl length of the double loss-of-function mutant (Fig. 4). Consistent with the results from our transgenic lines, our F1 hybrid data suggest that full ELF3 function depends on a permissive genetic background.

**Discussion**

Our results demonstrate that natural ELF3 polyQ variation that is not associated with disease has dramatic phenotypic consequences,
data are not consistent with intragenic suppressors. A loss-of-function mutant in the Ws background; elf3-col is the elf3-200 loss-of-function mutant in the Col background. Reciprocal crosses for each Fi showed similar results. n = 15–20 for each genotype, except for F1Ws × e1186Ws (n = 5). Error bars are SEM. Genotypes labeled with different letters differed significantly by Tukey-HSD test with α = 0.05.

Instead, ELF3-TR alleles seem “matched” to specific genetic backgrounds in which they are functional, whereas they are incompatible with other backgrounds. The haploinsufficiency of the elf3-col allele in Ws × Col hybrids supports this interpretation. In contrast, the ELF3-Ws allele is haploinsufficient in hybrids, indicating that the ELF3-Col × Ws incompatibility is asymmetric. This observation agrees with Orr’s assertion that incompatibility between recently diverged populations is usually asymmetrical, because it tends to arise from the derived allele (i.e., ELF3-Col) (43). Variable TRs, and the ELF3-TR in particular, have been previously suggested as agents of adaptation to new external environments (1, 16, 17, 20, 44). Our results suggest that polyQ-encoding TRs are also agents of coadaptation within genomes.

We speculate that the observed background effects arise from background-specific polymorphisms in genes encoding physically interacting proteins (26, 28–30). TRs have a far higher mutation rate than nonrepeated regions (10−4 per site per generation for TR vs. 10−8 for SNPs) (45, 46) and, as we show, their expansion or contraction can have dramatic phenotypic impact. ELF3’s partner proteins may have acquired compensatory mutations to accommodate new ELF3-TR variants and vice versa. Alternative explanations for the background effects are compensatory mutations in ELF3 (intragenic suppressors), or ELF3 interactions that are unique to a given background. Intragenic variation and protein modification can play an important role in polyQ-mediated phenotypes (47, 48). At least for the ELF3-Col allele, however, our F1 data are not consistent with intragenic suppressors.

Consistent with polyQ-mediated background effects, in at least one case a modifier mutation has been shown to delay onset of Huntington disease (11). Hypothetically, population genetic approaches could identify incompatible alleles that may contribute to variable disease onset in patients with polyQ expansions and to ELF3-dependent background effects in A. thaliana. However, the great diversity of TR alleles compared with SNP alleles and the small number of individuals carrying specific TR alleles render a population genetics approach infeasible. Extensive genetic mapping or other experimental approaches will be needed to identify the determinants of ELF3-TR-dependent background effects.

As TRs are rapidly evolving, we speculate that polyQ-mediated incompatibilities and the resulting fitness loss in hybrids and their offspring may contribute to disruption of gene flow between closely related populations. This speciation mechanism would be of particular importance for organisms with many polyQ-encoding TRs, thousands of offspring, and an inbreeding lifestyle. Even in humans, however, about 1% of proteins contain polyQ tracts (13, 14, 45). Our results identify TR copy number variation, and in particular polyQ variation, as a phenotypically important class of genetic variation that warrants genome-wide assessment in model organisms, crops, and humans alike.

Materials and Methods

Information on the genotypes and experimental growth conditions, recombinant DNA techniques and generation of ELF3 transgenic plants, and RNA extraction, processing and quantitative real-time PCR used in this study is provided in SI Materials and Methods. See Tables S5–S7 for primer information.

Developmental Phenotype Assays. Phenotypic assays were performed with expression-matched, homozygous T3 or T4 transgenic plants (Fig. S1 C and D and Tables S2–S4). For hypocotyl length and flowering-time experiments, plants were grown and measured in a pseudorandomized design under SD and LD conditions. More details are provided in SI Materials and Methods.

Luciferase Imaging and Period Analysis. Luciferase assays were performed with lines harboring the CCR2::LUC reporter. Further details are provided in SI Materials and Methods.

Principal Components Analysis. PCA on transformed phenotype data was performed using the R function prcomp (R Foundation for Statistical Computing, 2011, http://www.r-project.org/). More details are provided in SI Materials and Methods.

ACKNOWLEDGMENTS. We thank D. Nusinow for providing the ELF3 promoter, M. Nordborg for Arabidopsis thaliana accessions, and S. Fields for critical comments on the manuscript. This work was supported by National Human Genome Research Institute Interdisciplinary Training in Genome Sciences Grants T32HG000035-16 (to S.F.U.) and 2T32HG35-16 (to M.O.P.); Deutsche Forschungsgemeinschaft Grant DA1061/4-1 (to S.J.D.); and Royalty Fund Grant RRF4365 (to C.Q.).


