HSP90 buffers newly induced mutations in massively mutated plant lines

G. Alex Mason[a], Keisha D Carlson[b], Maximilian O Press[c], Kerry L Bubb[a#], Christine Queitsch[a#]

[a] Department of Genome Sciences, University of Washington, Seattle, WA, 98195, USA

[b] Department of Biology, University of Puget Sound, Tacoma, WA, 98416, USA

[c] Phase Genomics Inc., Seattle, WA, 98195

[#] Corresponding authors (kbubb@uw.edu; queitsch@uw.edu)

Short title: HSP90 buffers newly induced mutations in massively mutated plant lines

Abstract (300 words)

Robustness to both genetic and environmental change is an emergent feature of living systems. Loss of phenotypic robustness can be associated with increased penetrance of genetic variation. In model organisms and in humans, the phenotypic consequences of standing genetic variation can be buffered by the molecular chaperone HSP90. However, it has been argued that HSP90 has the opposite effect on newly introduced genetic variation. To test the buffering effect of HSP90 on new mutations, we introduced vast numbers of mutations into wild-type and HSP90-reduced plants and assessed embryonic lethality and early seedling phenotypes for thousands of offspring. Although the levels of newly introduced mutations were similar in the two backgrounds, the HSP90-reduced plants showed a significantly greater frequency of embryonic lethality and severe phenotypic abnormalities, consistent with higher penetrance and expressivity of newly introduced genetic variation. We further demonstrate that some mutant phenotypes were heritable in an HSP90-dependent manner, and we map candidate HSP90-dependent polymorphisms. Moreover, both sequence and phenotypic analyses of wild-type and HSP90-reduced plants suggest that the HSP90-dependent phenotypes are largely due the newly introduced mutations rather than to an increased mutation rate in HSP90-reduced plants. Taken together, our results support a model in which HSP90 buffers newly introduced mutations, and the phenotypic consequences of such mutations outweigh those of mutations arising *de novo* in response to HSP90 perturbation.

Introduction

In eukaryotes, the molecular chaperone HSP90 facilitates the folding of diverse protein clients (1). These clients are implicated in nearly all aspects of organismal growth and

development, and thus HSP90 is highly connected in genetic networks (1-3). As a heat shock protein and negative regulator of heat shock transcription factors (4), HSP90 is also linked to the external environment and enables environmental signaling through chaperoning receptors and downstream signaling components in many pathways (1, 3, 5, 6).

The interaction of HSP90 with its protein clients affords them additional physicochemical stability, allowing clients to accommodate mutations that may otherwise impact their function (7-11). Indeed, across the mammalian lineage, including humans, genes encoding HSP90 client kinases carry significantly more and more deleterious variation that those encoding non-client kinases (10). The higher evolutionary rate of HSP90 clients mirrors results with other chaperones (8, 9, 12-16), suggesting that protein stabilization generally increases evolutionary rates through relaxed selection on new mutations in clients (7).

Consequently, HSP90 inhibition compromises client protein folding and causes deleterious phenotypic effects, which may be exacerbated by genetic variation residing in the genes encoding these clients. This phenomenon is distinct from HSP90-dependent effects on epigenetic phenomena (17-19) or the frequency of *de novo* mutations (2, 20-25). HSP90 perturbation through mutation, pharmacological inhibition, or increased temperature reveals distinct, strain-specific morphological aberrations and altered quantitative traits in flies, plants, fish, and yeast (26-34). In worms, naturally varying HSP90 levels predict the penetrance of mutations in several signaling proteins (35, 36). In humans, HSP90 enables the function of several mutated oncogenic proteins, resulting in the chaperone serving as a drug target in cancer (6, 37-40).

HSP90 can affect both penetrance (*i.e.* frequency of an aberrant trait associated with a genetic variant) and expressivity (*i.e.* severity of an aberrant trait associated with a genetic

variant) (25, 29, 31, 33, 35, 36). The effect of HSP90 on variant penetrance and expressivity is the basis of the capacitor hypothesis which posits that HSP90 allows accumulation of genetic variation, which remains phenotypically silent until released upon stress or targeted HSP90 perturbation. However, two controversies, prompted by recent studies, remain to be resolved, the first concerning the nature of the variants that can be buffered by the chaperone (41), and the second concerning the relative contributions of HSP90-dependent *de novo* variation to phenotype (22, 24).

As to the first, the increased evolutionary rate of HSP90 clients across the mammalian lineage suggests that HSP90-buffered variation tends to be relatively old and remains present in populations in the face of selection (10). It is less clear how HSP90 affects new genetic variants or variant combinations that have not been exposed to selection. To address this question, studies in yeast and plants have studied hundreds of divergent segregant lines generated from crosses of outbred strains; these segregant lines represent new variant combinations that have not been exposed to long-term selection (28, 32, 33). In sum, the results argue that HSP90 buffers such new genetic combinations.

A recent study compared several yeast mutation accumulation lines, each carrying ~4 single nucleotide variants, to yeast strains isolated from natural environments to address whether HSP90 acts similarly on newly arising mutations and older standing variation (41). Focusing on trait variance rather than trait frequencies or trait means, Geiler-Samerotte and colleagues argue that HSP90 fails to confer mutational robustness for alleles that have not experienced selection. In contrast to the earlier studies (28, 32, 33), these authors also report diminished rather than increased phenotypic variance in response to HSP90 inhibition in yeast segregant lines, which they posit casts doubt on current strategies to target HSP90 in cancer therapies.

Thus far, no study has evaluated the effects of HSP90 on very large numbers of newly introduced mutations in an isogenic background, a strategy that should resolve these contradictions. Here, we introduced thousands of random mutations into Col-0 plants, the *Arabidopsis thaliana* reference genotype, and into Col-0 plants with reduced HSP90 levels (42) or treated with an HSP90 inhibitor (29). If HSP90 buffers at least some of the newly introduced mutations, we should, and did, observe a greater frequency and severity of mutant phenotypes in HSP90-reduced plants. The HSP90-responsive phenotypes in mutagenized seedlings commonly resembled those arising in the unmutagenized Col-0 background at much lower frequency (29), consistent with the notion that variants in a limited number of client proteins contribute to HSP90-dependent phenotypes as opposed to invoking complex higher-order epistasis across genetic networks. Broadly, our results support the idea that HSP90 buffers new genetic variation.

Lastly, our study design allowed us to address the controversy whether HSP90-dependent phenotypes are largely a consequence of *de novo* mutations arising in HSP90-reduced lines as recently argued (22). Although HSP90 inhibition increases the frequency of certain mutations (2, 20, 21, 23, 24, 43-48), we find no evidence that the majority of HSP90-dependent phenotypes observed here arises due to HSP90-dependent *de novo* mutations. In sum, our results resolve the two remaining controversies surrounding the HSP90 capacitor hypothesis.

Results

HSP90 perturbation increases penetrance of embryonic lethal mutations

We used ethyl methanesulfonate (EMS) mutagenesis to generate large numbers of new mutations in the *A. thaliana* reference strain Col-0 and in two Col-0-derived HSP90-RNAi lines, HSP90 RNAi-A1 and HSP90 RNAi-C1 (42) (**Figure 1A**). The construct in the HSP90 RNAi-C1 plants interferes with the function of multiple HSP90 paralogs to produce more severe HSP90-

dependent phenotypes than the construct in HSP90 RNAi-A1 plants, which interferes only with the heat-inducible HSP90.1 paralog (42).

The EMS-mutagenized seeds (M_1 generation) were transferred to soil and grown to maturity (**Figure 1A**). To assess the mutagenesis efficiency and the frequency of embryonic lethal mutations, we counted the number of live and dead seeds (M_2 generation) produced by the M_1 plants (**Figure 1B, C**) (49). As expected, higher concentrations of EMS led to significantly higher rates of embryonic lethal mutations relative to untreated Col-0 controls (**Figure 1C**). Values for seed viability and embryonic lethal mutations were reproducible across experiments. Untreated plants of the Col-0 and the two HSP90-reduced backgrounds produced similar numbers of live seeds (**Figure 1B**). In contrast, EMS-mutagenized HSP90 RNAi-C1 plants showed significantly more embryonic lethal mutations than EMS-mutagenized Col-0 plants (generalized linear model, family=Poisson: p<1.0E-06) (**Figure 1C, S1 Table**). However, EMS-mutagenized HSP90 RNAi-A1 plants did not show more embryonic lethal mutations than EMS-mutagenized Col-0 plants (**Figure 1C**).

A trivial explanation for the greater number of embryonic lethal mutations in the HSP90 RNAi-C1 background is either that these plants were more deeply mutagenized than Col-0 plants or that fewer mutations were repaired. HSP90 chaperones proteins that function in DNA repair (50, 51), and as HSP90 inhibition also increases the rate of certain mutations (2, 20-23, 43-48), even equivalent EMS treatments may produce more mutations in the HSP90 RNAi-C1 background. Alternatively, as predicted under the capacitor hypothesis (25, 31, 52-54), some EMS-generated mutations may acquire greater penetrance when HSP90 is inhibited and therefore result in embryonic lethality.

To distinguish whether increased penetrance or mutation quantity causes the increased embryonic lethality in the HSP90-reduced line, we performed whole-genome sequencing of six randomly selected EMS-mutagenized M_2 HSP90 RNAi-C1 and six randomly selected M_2 Col-0 plants. Both backgrounds showed similar genomic distributions and numbers of unique mutations per individual plant (**Figure 1D, E, S1 Fig**). Each M_2 individual showed ~1000 high-confidence, private single nucleotide variants (**Figure 1E, S1 Table,** x-coverage >= 5, derived allele coverage >= 2). We conclude that the significantly greater number of embryonic lethals observed in the EMS-mutagenized HSP90-reduced M_2 progeny are likely due to the increased penetrance of the newly introduced mutations because of reduced HSP90 buffering.

HSP90 perturbation increases the frequency of severely aberrant phenotypes across generations

Having established that newly introduced mutations are more likely to cause lethality in HSP90-reduced plants than in controls, we examined the effects of HSP90 reduction on the frequency of aberrant morphological phenotypes. As such plant phenotypes are qualitative in nature (29), we used a mutant scoring index to estimate the frequency of 16 complex early seedling traits, categorized by severity (**Figure 2A**, **B**).

We first examined the frequency of these aberrant phenotypes in the M_1 plants, in which the EMS-induced mutations are typically heterozygous and sometimes occur only somatically. Heterozygous mutations can show increased penetrance and cause dominant mutant phenotypes when HSP90 is inhibited (55). Indeed, M_1 HSP90 RNAi-C1 seedlings, but not M_1 HSP90 RNAi-A1 seedlings, displayed a modest yet significant increase in the frequency of individuals with mutant index scores ≥ 2 (Fisher's exact test: p-value<0.05, OR=1.5) (**S2 Fig**).

We used the highly specific HSP90 inhibitor geldanamycin (GdA) (Whitesell, 1994) to examine whether a similar trend held for the EMS-mutagenized Col-0 M_1 plants when HSP90 was inhibited pharmacologically. Indeed, GdA treatment significantly increased phenotype frequency and severity in this population (**S2 Fig**). We conclude that HSP90 perturbation increases the penetrance and expressivity of heterozygous or somatic mutations that affect seedling morphology phenotypes.

Second, we tested the response to HSP90 perturbation, both pharmacologically and genetically, in the M_2 generation using the same mutant phenotype index (**Figure 3A**). In the M_2 generation, with the newly introduced mutations segregating, some will be homozygous and others heterozygous. HSP90 perturbation should increase the expressivity and penetrance of these mutations and thereby cause a higher frequency of severely aberrant morphological phenotypes. Consistent with this expectation, we observed a significantly greater frequency of severe phenotypes among the EMS-mutagenized seedlings grown with the HSP90 inhibitor compared to unmutagenized seedlings and mutagenized seedlings grown with the DMSO solvent (**Figure 3B**). This observation manifests statistically as an interaction between GdA treatment and mutagenesis in their effect on phenotype frequency; the strongest such interaction was observed at an intermediate GdA dose (Fisher's exact test: p-value <1.0E-15, OR=4.6). Genetic HSP90 perturbation yielded similar results: the M_2 HSP90 RNAi-C1 seedlings produced significantly more aberrant phenotypes than comparable M_2 Col-0 seedlings (Fisher's exact test: p-value <1.0E-15, OR=8.8) (**Figure 3C**). We conclude that HSP90 perturbation increases the frequency and severity of aberrant phenotypes after mutagenesis.

Tracking the HSP90-dependent nature of aberrant morphological phenotypes across generations

We asked whether the phenotypes from newly introduced mutations are indeed HSP90-dependent, *i.e.* whether or not they would be expressed when HSP90 is fully functional. As it is not possible to examine a single plant both with and without HSP90 perturbation, we generated M_3 offspring from M_2 parents that displayed aberrant phenotypes under HSP90-reduced or control conditions (**see Materials and Methods**). Although some of the newly generated EMS mutations will segregate in the M_3 offspring, this approach generates groups of sibling plants with a similar mutation background that can be examined under either control or HSP90-reduced conditions. Specifically, we selected M_2 plants with the deformed aerial organ phenotype on DMSO medium (DM plants, Figure 4 A, B) or GdA medium (GM plants). Deformed aerial organs are readily identified and can be a consequence of HSP90 perturbation (29, 42). In addition to the deformed aerial organ phenotype, we also examined the M_3 progeny of the 22 GM and 10 DM plants for the other 15 phenotypes (**Figure 2A, B**) because more than one phenotype was noted for many parental M_2 plants (**S3 Table).**

Under standard growth conditions, 20 M_3 lines showed significantly increased frequencies of deformed aerial organs, with more DM lines (7 out of 10) than GM lines (13 out of 22) producing this phenotype (**Figure 4A, B**). The lesser heritability of this trait in the GM lines is consistent with HSP90 perturbation increasing penetrance and expressivity of newly introduced variation in the parental lines.

Next, we asked how many of these lines showed an HSP90-dependent response in the frequency of this phenotype. Nine of 13 GM lines showed significant HSP90-dependence in the expression of the aberrant phenotype, whereas only one out of seven DM lines did. This DM line

and seven out of the nine significantly HSP90-responsive GM lines showed low frequencies of deformed aerial organs in control conditions (*i.e.* DMSO) and significantly increased phenotype frequencies when HSP90 was perturbed (**Figure 4C**, GM, dark red lines, DM, bright red line). This HSP90-dependent increase in phenotype frequencies agrees with several previous studies (29, 30, 34, 36, 56). We also found two GM lines with a significantly decreased frequency of the aberrant phenotype (**Figure 4C**, GM dark blue lines, Fisher's exact test: p-adj<0.05, OR<0.03), reminiscent of previous findings that HSP90 perturbation can conceal the phenotypic consequences of underlying variation (26, 28, 32, 41).

As some parental M_2 plants showed other aberrant phenotypes in addition to the selected phenotype, it was not surprising to find these phenotypes also in their offspring (**Figure 4B**). Only two DM M_2 plants with deformed aerial organs showed additional defects; both produced M_3 offspring with these defects; none were HSP90-responsive. In stark contrast, 19 GM M_2 showed additional aberrant phenotypes; 16 of these produced M_3 offspring with these defects; significant response to HSP90 perturbation was observed in nine M_3 lines **(S3 Fig)**.

Assessing all aberrant phenotypes together in the context of genetic background and response to HSP90 perturbation (**Figure 4D**; **S3 Fig; S4 Fig**) revealed important trends. First, consistent with previous data (29, 56), upon HSP90-perturbation, unmutagenized Col-0 wild-type plants tended to show dwarf phenotypes, lesions on cotyledons, discoloration, and epinastic organs among others. Second, about half of the mutagenized M_3 also showed these phenotypes in response to HSP90 perturbation, albeit often at far higher frequency. Assuming that the HSP90-dependent phenotypes observed in wild-type Col-0 plants reflect dysfunctional client proteins (25, 57), this observation suggests that some of the newly introduced HSP90-dependent variants reside in genes encoding these clients or network nodes closely associated with them. Third,

selection for the deformed aerial organ trait in the mutagenized lines increased the phenotype frequencies dramatically in some lines, consistent with a genetic, possibly multigenic basis. Lastly, certain traits, such as aborted seedlings and severe developmental delay that are observed only infrequently in unmutagenized Col-0, were strongly responsive to HSP90 perturbation in the mutagenized M_3 lines, indicative of problems in early plant development and consistent with the increased embryonic lethality observed in the M_1 generation. Taken together, our results indicate that at least some newly introduced mutations readily lead to heritable HSP90-responsive phenotypes.

Mapping HSP90-responsive loci containing newly introduced mutations

To further explore the genetic basis of the HSP90-responsive deformed aerial organ trait, we applied a bulk segregant approach, notwithstanding the significant challenges due to the large number of introduced EMS mutations and the partially penetrant, HSP90-responsive nature of the trait. As the aberrant phenotype was both partially penetrant in the mutagenized lines and present at low levels in Col-0 wild-type seedlings, we would not expect a derived allele frequency of 100% indicating the approximate location of the causal mutation, as expected in a traditional bulk segregant analysis (58). Rather, we expected a contiguous increase in derived allele frequency (DAF) above background levels expected in an F_2 population in Hardy-Weinberg equilibrium (DAF=0.5).

We selected two M_3 lines, GM6 and GM9, with significantly increased penetrance of this phenotype under HSP90-reduced conditions (OR>4.5). Single GM6 and GM9 M_3 individuals with deformed aerial organs were backcrossed to Col-0 wild-type plants to generate heterozygous F_1 populations. A randomly selected individual from each cross was selfed to generate segregating F_2 populations for both backgrounds (**Figure 5A, B; S5 Fig**). As this

backcross will dilute phenotype-associated variation, fewer seedling with deformed aerial organs were observed in standard growth conditions while phenotype frequencies in HSP90-reduced conditions stayed high (**Figure 5C**; **S5 Fig**). To map the HSP90-responsive EMS mutations in the GM6 and GM9 backgrounds, we sequenced pools of approximately 120 F_2 individuals displaying deformed aerial organs upon HSP90 perturbation. For both backgrounds, we found regions with elevated derived allele frequency, consistent with the presence of newly introduced HSP90-responsive mutations (**Figure 5D; S5 Fig; S6 Fig**).

For the GM6 F_2 population, one notable region contained eight consecutive new mutations with elevated derived allele frequency. Six of these did not reside in coding regions, but one (Chr1:5,407,084 C=>T) was a missense mutation (proline to leucine) in the MYB family transcription factor AT1G15720 (**Figure 5D, S4 Table**). For the GM9 F_2 population, one region contained six consecutive new mutations with elevated derived allele frequency. Five of these resided in coding regions, each in different genes, and each was a missense mutation (**S5 Fig**), two of which had negative blosum62 scores, suggesting deleterious effects (**S4 Table**).

To determine how likely these observations were to occur by chance (*i.e.* no new mutation conferred increased phenotypic penetrance upon HSP90-reduction), we performed simulations in which 240 gametes (representing 120 random F_2 individuals, carrying two haploid pseudochromosomes each) were generated from a simulated F_1 individual, heterozygous for n mutations, assuming one recombination event per chromosome. Of these 240 pseudochromosomes, we sampled x, reflecting actual experimental x-coverage (read depth, in GM6 mean x-coverage =77; in GM9 mean x-coverage =54). Out of 100 such simulations, none resulted in 5 or more consecutive SNVs with a derived allele frequency greater than 0.65. (**S7 Fig**). Thus, despite the challenges of mapping the genetic determinants of a likely multigenic,

partially penetrant trait, our results provide additional evidence for the existence of newly introduced HSP90-buffered variants and support for candidate loci containing them.

Comparison of the derived allele frequencies in unmutagenized Col-0 wild-type and HSP90-reduced lines

Our study design and sequencing analyses lend themselves to addressing another controversy surrounding the capacitor hypothesis (22): are HSP90-dependent phenotypes largely a consequence of increased frequency of *de novo* germline mutations in response to HSP90 perturbation?

Our Col-0 wild-type (provided in 1994 by Brian Keith) and the HSP90 RNAi lines derived from it have been bulk-propagated in the laboratory for two decades. Therefore, these backgrounds are expected to differ in genome sequence from the *A. thaliana* reference genome. A mutation associated with altered response to HSP90 perturbation would result in an increase in the derived allele frequency for an entire region (haplotype block) containing multiple SNVs, as we describe above for GM6 and GM9. Therefore, we were conservative in annotating SNVs as likely EMS-derived and removed any SNVs that were present pre-mutagenesis in our Col-0 and HSP90-reduced lines compared to the reference genome. To do so, we generated sequence data for Col-0 and the HSP90-RNAi C1 line, examining the number of derived alleles (compared to the *A. thaliana* reference genome) in single plants of both backgrounds (two Col-0 plants, two HSP90 RNAi-C1 plants). As single individuals were sequenced to ~10x coverage, we counted bases only with derived allele frequency of 1 (homozygous for derived allele). We found similar numbers of derived alleles in all four plants (**S1 Table**). There was no evidence supporting higher numbers of single nucleotide mutations in HSP90 RNAi-C1 compared to wild-type Col-0.

Although HSP90 perturbation can lead to *de novo* germline (and likely somatic) mutations, it appears unlikely that these mutations contributed significantly to the greater penetrance and expressivity of aberrant phenotypes in the HSP90-reduced seedlings. This assertion is further supported by the fact that the unmutagenized HSP90 RNAi lines showed similar levels of embryonic lethality as Col-0 wild-type (**Figure 1 B, C**). Moreover, we reproduced the results generated with the HSP90 RNAi-C1 line with pharmacological inhibition of HSP90 in mutagenized Col-0 seedlings. The HSP90 inhibitor is supplied to the growth medium and taken up during seed germination and growth; the seeds contain fully formed embryos. An HSP90-dependent *de novo* mutation affecting these embryos would be somatic and hence unlikely to be heritable. Taken together, our results support the argument that the increased frequency and severity of aberrant phenotypes in response to HSP90 perturbation is due to increased penetrance and expressivity of newly introduced EMS mutations rather than germline *de novo* mutations arising in the HSP90 RNAi-C1 line.

Discussion

Our data argue that HSP90 buffers the penetrance and expressivity of newly introduced genetic variation in *A. thaliana*. Penetrance effects included an increased frequency of embryonic lethals, and expressivity effects included more severe aberrant morphological phenotypes. Tracking phenotypes across generations, we found several lines that showed heritable aberrant morphological phenotypes in an HSP90-dependent manner. As it does for standing genetic variation (25, 28-31, 34, 59-61), HSP90 plays a role in buffering phenotype against newly introduced genetic variation, which has not been exposed to selection.

These observations are broadly consistent with past studies on new genetic combinations in yeast and plants (28, 32, 33). However, one recent yeast study (41) concludes that phenotypic variance associated with newly introduced mutations typically decreases under HSP90 inhibition (*i.e.* "potentiation", 25, 26, 62). Here, we argue the opposite: the more parsimonious interpretation of the available data is that HSP90 commonly buffers new mutations.

Our argument relies on multiple lines of evidence. First, we screened a large number of mutations. The mutagenized and sequenced M_2 individuals carried between 800 to 2,000 new mutations, such that we screened on the order of tens of millions of new mutations, many of them through multiple generations. In contrast, the 94 mutation accumulation lines tested in the recent yeast study carry on average only four single nucleotide mutations per haploid genome (41). Second, we demonstrated the substantial heritability of HSP90-responsive traits, suggesting that these phenotypes are likely to be the effect of genetic mutations rather than stochastic effects. Third, we identified candidate loci associated with an HSP90-responsive phenotype in two independent lines. Fourth, we used two methods (genetic and pharmacological) to perturb HSP90, whose results corroborated each other, demonstrating that the findings are not a consequence of either geldanamycin's off-target effects or a peculiarity of the HSP90 RNAi lines. Fifth, all mutations were assayed immediately after introduction into the *A. thaliana* genome, such that our analysis should represent the full spectrum of phenotypic effects of new mutations, including those that would be quickly purged by selection. Last, in prior studies, we demonstrate that the effects of revealed HSP90-responsive variation on phenotype means outweigh its effects on phenotype variance by as much as a magnitude, consistent with the capacitor hypothesis' premise that HSP90-dependent phenotypes can be acted upon by natural selection (33).

However, the present study has some limitations. First, the semi-quantitative phenotype categories make estimates of heritability subject to threshold effects, likely decreasing heritability estimates. Second, aside from embryo lethality, we assayed highly complex developmental phenotypes, and the effects of HSP90 on these phenotypes may be less relevant in the unicellular yeast, which does not experience a prolonged ontogeny. Third, EMS mutagenesis may introduce a different spectrum of mutations than those that arise spontaneously in natural populations.

That HSP90 can buffer new mutations is also supported by newly emerging data. For example, saturation mutagenesis of the non-client yeast transcription factor Ste12 revealed HSP90-responsive Ste12 variants; these variants were rare and position-dependent; moreover, they were implicated in the environmentally-sensitive balance between the mutually exclusive traits of yeast mating and invasion that are governed by Ste12 (63). Similarly, recurrent HSP90-dependent variants in oncogenes are highly sequence-specific and rare (60, 64-66). Notwithstanding these results on the buffering of new mutations, standing variation is likely a more significant actor in most HSP90-dependent phenotypes (11, 25-34, 56, 59).

Materials and Methods

Plant Growth Conditions: Seedlings of the Columbia-0 (Col-0) ecotype were used as a wild-type reference; this stock has been propagated in the Lindquist and Queitsch labs since 1994 and was originally provided by Brian Keith. The HSP90 RNAi-A1 and HSP90 RNAi-C1 were generated in this background (42). For experiments, seeds were sterilized with ethanol and plated onto 1× Murashige and Skoog (MS) basal salt medium supplemented with 1× MS vitamins, 1% (wt/vol) sucrose, 0.05% MES (wt/vol), and 0.24% (wt/vol) Phytagel. After

stratification in the dark at 4 °C for 5 days, plates were transferred to an incubator

(Conviron) set to long days (LD) (16L:8D at 22 °C:20 °C), with light supplied at 100

μmol·m-2·s-1 by cool-white fluorescent bulbs; for geldanamycin experiments, plates were

placed under long-pass yellow filters that blocks 454 nm light (67). Seedlings were scored for

mutant phenotypes at day 10. Geldanamycin was purchased from Sigma Aldrich (G3381),

diluted in DMSO, and used for M_1 and M_2 experiments. As this reagent was no longer available

from Sigma at reasonable cost, another lot of geldanamycin was purchased from LC Laboratories

(G-4500), diluted in DMSO, and used in the M_3 heritability experiments. This reagent showed

stronger phenotypic effects than Sigma geldanamycin and hence was used at a lower dose (0.5

μM instead of 0.6 μM)

EMS Mutagenesis: For each genotype and treatment, 50 mg of seeds were weighed out (~2500

seeds). Seeds were sterilized by a ten-minute wash in 70% ethanol (v/v) with 0.1% Triton-X

(v/v), followed by a five-minute wash with 95% ethanol (v/v) and one wash with sterile water.

Seeds were then transferred to a tube containing 12 ml of 0.0005% Tween (v/v). EMS (ethyl

methanesulfonate, Sigma-Aldrich, M0880) was added to a final concentration of either 0.2%

(v/v) or 0.4% (v/v) in each tube. Control samples received only Tween treatment without EMS.

Tubes were rotated at room temperature for 15 hours. Seeds were washed eight times for ten

minutes in sterile water and then plated on indicated plant growth medium or sown on soil.

Determining viable seeds and embryonic lethality in M_1 progeny: After mutagenesis, M_1

seeds were sown on soil and grown as indicated. For each genotype and treatment, one silique

was collected from ~ 50 adult plants. For each silique, the number of viable and dead seeds was

counted. Dead seeds are representative of individuals with lethal mutations and were used to estimate mutation burden after mutagenesis (49). We calculated the average number of embryonic lethal mutations per diploid genome (μ) as follows: if X equals the number of siliques without dead seeds, and Y equal the total number siliques observed, $\mu = -\ln(X/Y)$. If μ is greater than one (64% or more of observed siliques have dead seeds; $1-X/Y = 0.64$), mutagenesis is successfully saturated (49).

Assessing the frequency of aberrant morphological phenotypes in the M_1 and M_2 generation: We devised a mutant scoring index to estimate frequency of mutant-like phenotypes quickly and objectively after mutagenesis. We based this index on known HSP90-responsive phenotypes, such as size-related traits or lesions on aerial tissue (29, 42, 56). In this index, 16 complex early seedling traits are categorized by severity (**Figure 2A**). Traits such severe developmental delay and lethality were weighted as more deleterious. For example, a seedling with the phenotype of severe developmental delay received a score of 3, while a seedling with epinastic cotyledons received a score of 1 (**See Figure 2B for examples**). The mutant scoring index also accounted for seedlings with multiple mutant phenotypes; these received higher index scores (**See Figure 2B for examples**).

Seedlings were grown as previously described (29, 42). Seedlings of indicated genotypes and treatments were scored blindly for the described morphological phenotypes (**Figure 2**). Seedlings were photographed with a Canon Power Shot S5 IS camera, pictures were edited using Adobe Photoshop CC 2017. Edits addressed image size, brightness, and contrast and did not alter depiction of seedling phenotypes.

M_3 heritability experiments: 65 M_2 individuals with aberrant, deformed aerial organ phenotypes, grown on plates containing either 0.5 µM GdA or DMSO, were selected, transferred to soil, and propagated to generate M_3 seeds (see **S3 Table for parental phenotypes**). For the selected seedlings, ~50% survived the transfer and produced progeny. The resulting M_3 progeny, designated either DM (progeny of M_2 plant with selected phenotype grown on DMSO medium) or GM (progeny of M_2 plant with selected phenotype grown on GdA medium), was assayed for 16 early-seedling phenotypes in the presence of either 0.5 µM GdA or DMSO (GdA solvent, mock treatment). For phenotyping, M_3 lines were grown as indicated and scored blindly for the 16 morphological phenotypes (**Figure 2**); blind scoring relied on scrambled line identifiers. Seedlings were photographed with a Canon Power Shot S5 IS camera and were edited using Adobe Photoshop CC 2017.

A given trait in a given M_3 line was considered responsive to HSP90 perturbation if $X_{GdA, M3} - X_{DMSO, M3} \geq 10\%$ (positive) or if $X_{DMSO, M3} - X_{GdA, M3} \leq 10\%$ (negative) with $X = (\%)$ phenotype frequency of affected seedlings. A trait was considered unresponsive to HSP90 perturbation if $X_{GdA, M3} - X_{DMSO, M3} \approx 0$. Further, to account for effects of HSP90 perturbation on traits commonly observed in the wild-type Col-0 background, we considered traits with $X_{GdA, M3} - X_{DMSO, M3} \approx X_{GdA, \text{wild-type Col-0}} - X_{DMSO, \text{wild-type Col-0}}$ as showing no effect of newly introduced EMS mutations. Comparisons of seedling phenotype frequencies were performed with Fisher's exact test to determine statistical significance.

SNV calling on two unmutagenized Col-0, two unmutagenized HSP90 RNAi-C1, six mutagenized Col-0 and six mutagenized HSP90 RNAi-C1 individual plants: Genomic DNA was extracted from one adult leaf per individual using the CTAB method. Total DNA was

quantified with Qubit HS dsDNA assays according to the manufacturer's instructions. Sequencing libraries were generated using 50 ng of input DNA using the Nextera (Illumina) sample kit according to the manufacturer's instructions. Libraries were quality checked on the Agilent 2100 bioanalyzer using a DNA high-sensitivity chip (Agilent). The samples were sequenced by Genewiz on an Illumina HiSeq (100bp paired-end reads).

WGS reads were aligned to the TAIR10 release genome (http://www.arabidopsis.org/) using the default parameters of BWA (68). Aligned reads with mapping quality less than 35 were discarded. A VCF file was generated using SAMtools (http://samtools.sourceforge.net). Variants were further filtered for minimum coverage = 5 and minimum derived allele coverage = 2. We filtered each VCF file for unique mutations (*i.e.* unshared with any other sample within either Col-0 or HSP90 RNAi-C1 genotypes) as these were likely due to the EMS treatment. After identifying unique and private SNVs, we determined the number of total non-reference transitions and transversions for each individual. For analysis in mutagenized backgrounds, a similar filtering process was performed, where again we determined unique, unshared non-reference variants in each sequenced individual. See **S1 Table** for more information about the called variants.

Bulk segregant sequencing analysis: We backcrossed M_3 individuals with deformed aerial organs from the GM6 and GM9 lines to Col-0 wild-type plants used for mutagenesis. The resulting F_1 plants were propagated to the F_2 generation. For each backcross, F_2 seedlings were grown on HSP90-reduced growth medium (0.5 µM GdA), seedlings with deformed aerial organs were selected and pooled (n>90). Total DNA was extracted from these pools using the CTAB

method; DNA was quantified using the Qubit HS dsDNA assayed. Sequencing libraries were

generated using 10 ng input DNA with the Nextera (Illumina) sample kit. Libraries were quality

checked on the Agilent 2100 bioanalyzer using a DNA high-sensitivity chip (Agilent). The

samples were sequenced using an Illumina NextSeq in 150bp paired-end run.

Because our Col-0 line has been bulk-propagated in our lab since 1994, wild-type Col-0

individuals contain derived SNVs with respect to the published TAIR10 reference sequence at

varying frequencies. We identified these background SNVs by sequencing two pooled Col-0

individuals at ~8 x-coverage, allowing us to remove such derived SNVs from our list of EMS-

generated mutations. Additional background Col-0 SNVs were identified by comparing SNVs in

the deeply sequenced (>50 x-coverage each) F_2 populations derived from GM6 and GM9.

Derived SNVs found in both populations are unlikely to be due to EMS treatment.

With the pooled F_2 plants, we are measuring derived allele frequency at likely EMS-generated

SNVs, a task that requires much greater read depth. Furthermore, SNVs present in the

unmutagenized parents (M_0) relative to TAIR10 reference genome must be carefully filtered out.

As we are attempting to identify regions rather than single SNVs with elevated derived allele

frequency, erroneously discarding or including certain SNVs should not prevent the

identification of regions containing causative SNVs.

Our filtering pipeline proceeded as follows:

First, reads from the F_2 pools were mapped to the TAIR10 reference genome using BWA

(default parameters). Mapped reads outside of the centromeres (chr1:13,698,788-15,897,560;

chr2:2,450,003-5,500,000; chr3:11,298,763-14,289,014; chr4:1,800,002-5,150,000;

chr5:10,999,996-13,332,770) with map quality > 20 were maintained (69).

Next, SNVs were identified in a pool of M_0 Col-0 seedlings to identify background SNVs as well

as in the two pools of F_2 seedlings (aberrant seedlings grown on GDA-reducing media, for both

the GM6 and GM9 strains). All bases with any evidence of an alternate allele were identified

(derived allele frequency > 0). We excluded background SNVs (DAF>0) identified (i) in the

Col-0 pool (n=1,022,369), and (2) in both the two deeply sequenced GM6 and GM9 GDA-

treated F_2 pools (n=1,093,093). There was very little overlap between these sets of background

SNVs (n=78,805), indicating that the vast majority of the former set were either spuriously

identified as polymorphic -- which is certainly possible given the low sequencing coverage -- or

neither of the individual seeds that were mutated to give rise to the GM6 and GM9 lines

contained the derived allele -- also possible, given that many of these flagged bases had

relatively low derived allele frequency.

Among this filtered set of EMS-derived SNVs identified in the two F_2 pools, we then attempted

to remove SNVs erroneously arising due to sequencing error by removing SNVs with less than

two reads indicating the derived allele. Reads collected typically had per-bp Phred quality score

of 20 or more (error rate < 1/1,000), therefore the probability of two reads presenting an identical

error is low.

We also attempted to remove SNVs erroneously arising due to mapping error by removing SNVs

with unusually low x-coverage. The mean x-coverage in the GDA-treated F_2 pools was high

(GM6 F_2 x-coverage=77; GM9 F_2 x-coverage=54) and followed an approximately lognormal distribution, per base. Specifically, we required each such bp to have xcov greater than mean(log10(x-coverage)) - 2*sd(log10(x-coverage)).

For our final analysis, we examined only likely EMS mutations with derived allele frequency > 0.2.

Statistical analysis: All statistical analyses were performed in R Version 3.2.5. Comparisons of seedling aberrant phenotype frequencies were performed with Fisher's exact test (**S2 Table**). Comparisons of embryonic lethality were modeled with a Poisson regression using a general linear model. Student's t-tests performed for seed viability. When appropriate, p-values were adjusted using the R function "p.adjust," with method= "fdr." The results of all statistical tests performed are in **S2 Table**.

Acknowledgments

References

1.   Schopf FH, Biebl MM, Buchner J. The HSP90 chaperone machinery. Nature reviews Molecular cell biology. 2017 Jun;18(6):345-60.

2.   Queitsch C, Carlson KD, Girirajan S. Lessons from model organisms: phenotypic robustness and missing heritability in complex disease. PLoS genetics. 2012;8(11):e1003041.

3.   Taipale M, Jarosz DF, Lindquist S. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. Nat Rev Mol Cell Biol. 2010 Jul;11(7):515-28.

4.   Zou J, Guo Y, Guettouche T, Smith DF, Voellmy R. Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. Cell. 1998 Aug 21;94(4):471-80.

5.   Sangster TA, Queitsch C. The HSP90 chaperone complex, an emerging force in plant development and phenotypic plasticity. Curr Opin Plant Biol. 2005 Feb;8(1):86-92.

6.   Whitesell L, Lindquist SL. HSP90 and the chaperoning of cancer. Nat Rev Cancer. 2005 Oct;5(10):761-72.

7.   Bloom JD, Labthavikul ST, Otey CR, Arnold FH. Protein stability promotes evolvability. Proc Natl Acad Sci U S A. 2006 Apr 11;103(15):5869-74.

8.   Bogumil D, Dagan T. Cumulative impact of chaperone-mediated folding on genome evolution. Biochemistry. 2012;51(50):9941-53.

9.   Bogumil D, Landan G, Ilhan J, Dagan T. Chaperones divide yeast proteins into classes of expression level and evolutionary rate. Genome biology and evolution. 2012;4(5):618-25.

10.   Lachowiec J, Lemus T, Borenstein E, Queitsch C. Hsp90 promotes kinase evolution. Mol Biol Evol. 2015 Jan;32(1):91-9.

11.   Lachowiec J, Lemus T, Thomas JH, Murphy PJ, Nemhauser JL, Queitsch C. The protein chaperone HSP90 can facilitate the divergence of gene duplicates. Genetics. 2013 Apr;193(4):1269-77.

12.   Aguilar-Rodriguez J, Sabater-Munoz B, Montagud-Martinez R, Berlanga V, Alvarez-Ponce D, Wagner A, et al. The Molecular Chaperone DnaK Is a Source of Mutational Robustness. Genome biology and evolution. 2016 Oct 5;8(9):2979-91.

13.   Cetinbas M, Shakhnovich EI. Catalysis of protein folding by chaperones accelerates evolutionary dynamics in adapting cell populations. PLoS Comput Biol. 2013;9(11):e1003269.

14.   Sabater-Munoz B, Prats-Escriche M, Montagud-Martinez R, Lopez-Cerdan A, Toft C, Aguilar-Rodriguez J, et al. Fitness Trade-Offs Determine the Role of the Molecular Chaperonin GroEL in Buffering Mutations. Mol Biol Evol. 2015 Oct;32(10):2681-93.

15.   Tokuriki N, Tawfik DS. Chaperonin overexpression promotes genetic variation and enzyme evolution. Nature. 2009 Jun 4;459(7247):668-73.

16.   Williams TA, Fares MA. The effect of chaperonin buffering on protein evolution. Genome Biol Evol. 2010;2:609-19.

17.   Ruden DM, Lu X. Hsp90 affecting chromatin remodeling might explain transgenerational epigenetic inheritance in Drosophila. Curr Genomics. 2008 Nov;9(7):500-8.

18.   Sollars V, Lu X, Xiao L, Wang X, Garfinkel MD, Ruden DM. Evidence for an epigenetic mechanism by which Hsp90 acts as a capacitor for morphological evolution. Nature genetics. 2003 Jan;33(1):70-4.

19.   Tariq M, Nussbaumer U, Chen Y, Beisel C, Paro R. Trithorax requires Hsp90 for maintenance of active chromatin at sites of gene expression. Proc Natl Acad Sci U S A. 2009 Jan 27;106(4):1157-62.

20.   Chen G, Bradford WD, Seidel CW, Li R. Hsp90 stress potentiates rapid cellular adaptation through induction of aneuploidy. Nature. 2012 Feb 9;482(7384):246-50.

21.   Dote H, Burgan WE, Camphausen K, Tofilon PJ. Inhibition of hsp90 compromises the DNA damage response to radiation. Cancer Res. 2006 Sep 15;66(18):9211-20.
22.   Fanti L, Piacentini L, Cappucci U, Casale AM, Pimpinelli S. Canalization by Selection of de Novo Induced Mutations. Genetics. 2017 Aug;206(4):1995-2006.
23.   Gangaraju VK, Yin H, Weiner MM, Wang J, Huang XA, Lin H. Drosophila Piwi functions in Hsp90-mediated suppression of phenotypic variation. Nat Genet. 2011 Feb;43(2):153-8.
24.   Specchia V, Piacentini L, Tritto P, Fanti L, D'Alessandro R, Palumbo G, et al. Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. Nature. 2010 Feb 4;463(7281):662-5.
25.   Zabinsky RA, Mason GA, Queitsch C, Jarosz DF. It's not magic - Hsp90 and its effects on genetic and epigenetic variation. Semin Cell Dev Biol. [Review]. 2018 May 25.
26.   Cowen LE, Lindquist S. Hsp90 potentiates the rapid evolution of new traits: drug resistance in diverse fungi. Science. 2005 Sep 30;309(5744):2185-9.

27.   Hummel B, Hansen EC, Yoveva A, Aprile-Garcia F, Hussong R, Sawarkar R. The evolutionary capacitor HSP90 buffers the regulatory effects of mammalian endogenous retroviruses. Nat Struct Mol Biol. 2017 Mar;24(3):234-42.
28.   Jarosz DF, Lindquist S. Hsp90 and environmental stress transform the adaptive value of natural genetic variation. Science. 2010 Dec 24;330(6012):1820-4.
29.   Queitsch C, Sangster TA, Lindquist S. Hsp90 as a capacitor of phenotypic variation. Nature. 2002 Jun 6;417(6889):618-24.
30.   Rohner N, Jarosz DF, Kowalko JE, Yoshizawa M, Jeffery WR, Borowsky RL, et al. Cryptic variation in morphological evolution: HSP90 as a capacitor for loss of eyes in cavefish. Science. 2013 Dec 13;342(6164):1372-5.
31.   Rutherford SL, Lindquist S. Hsp90 as a capacitor for morphological evolution. Nature. 1998 Nov 26;396(6709):336-42.
32.   Sangster TA, Salathia N, Lee HN, Watanabe E, Schellenberg K, Morneau K, et al. HSP90-buffered genetic variation is common in Arabidopsis thaliana. Proc Natl Acad Sci U S A. 2008 Feb 26;105(8):2969-74.
33.   Sangster TA, Salathia N, Undurraga S, Milo R, Schellenberg K, Lindquist S, et al. HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. Proc Natl Acad Sci U S A. 2008 Feb 26;105(8):2963-8.
34.   Yeyati PL, Bancewicz RM, Maule J, van Heyningen V. Hsp90 selectively modulates phenotype in vertebrate development. PLoS Genet. 2007 Mar 30;3(3):e43.
35.   Burga A, Casanueva MO, Lehner B. Predicting mutation outcome from early stochastic variation in genetic interaction partners. Nature. 2011 Dec 8;480(7376):250-3.
36.   Casanueva MO, Burga A, Lehner B. Fitness trade-offs and environmentally induced mutation buffering in isogenic C. elegans. Science. 2012 Jan 6;335(6064):82-5.

37.   Calderwood SK, Neckers L. Hsp90 in Cancer: Transcriptional Roles in the Nucleus. Adv Cancer Res. 2016;129:89-106.

38.   Neckers L. Heat shock protein 90: the cancer chaperone. Journal of biosciences. [Review]. 2007 Apr;32(3):517-30.
39.   Trepel J, Mollapour M, Giaccone G, Neckers L. Targeting the dynamic HSP90 complex in cancer. Nature reviews Cancer. [Review]. 2010 Aug;10(8):537-49.

40.   Whitesell L, Santagata S, Mendillo ML, Lin NU, Proia DA, Lindquist S. HSP90 empowers evolution of resistance to hormonal therapy in human breast cancer models. Proceedings of the National Academy of Sciences of the United States of America. 2014 Dec 23;111(51):18297-302.

41.   Geiler-Samerotte KA, Zhu YO, Goulet BE, Hall DW, Siegal ML. Selection Transforms the Landscape of Genetic Variation Interacting with Hsp90. PLoS Biol. 2016 Oct;14(10):e2000465.

42.   Sangster TA, Bahrami A, Wilczek A, Watanabe E, Schellenberg K, McLellan C, et al. Phenotypic diversity and altered environmental plasticity in Arabidopsis thaliana with reduced Hsp90 levels. PLoS ONE. 2007;2(7):e648.

43.   Ichiyanagi T, Ichiyanagi K, Ogawa A, Kuramochi-Miyagawa S, Nakano T, Chuma S, et al. HSP90alpha plays an important role in piRNA biogenesis and retrotransposon repression in mouse. Nucleic Acids Res. 2014 Oct 29;42(19):11903-11.

44.   Khurana N, Laskar S, Bhattacharyya MK, Bhattacharyya S. Hsp90 induces increased genomic instability toward DNA-damaging agents by tuning down RAD53 transcription. Mol Biol Cell. 2016 Aug 1;27(15):2463-78.

45.   Kozeko L, Talalaiev O, Neimash V, Povarchuk V. A protective role of HSP90 chaperone in gamma-irradiated Arabidopsis thaliana seeds. Life Sci Space Res (Amst). 2015 Jul;6:51-8.

46.   Mittelman D, Sykoudis K, Hersh M, Lin Y, Wilson JH. Hsp90 modulates CAG repeat instability in human cells. Cell Stress Chaperones. 2010 Sep;15(5):753-9.

47.   Ryan CP, Brownlie JC, Whyard S. Hsp90 and Physiological Stress Are Linked to Autonomous Transposon Mobility and Heritable Genetic Change in Nematodes. Genome biology and evolution. 2016 Dec 1;8(12):3794-805.

48.   Xiol J, Cora E, Koglgruber R, Chuma S, Subramanian S, Hosokawa M, et al. A role for Fkbp6 and the chaperone machinery in piRNA amplification and transposon silencing. Mol Cell. 2012 Sep 28;47(6):970-9.

49.   Weigel D, Glazebrook J, editors. Arabidopsis: A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2002.

50.   Echeverria PC, Bernthaler A, Dupuis P, Mayer B, Picard D. An interaction network predicted from public data as a discovery tool: application to the Hsp90 molecular chaperone machine. PLoS ONE. 2011;6(10):e26044.

51.   Picard D. https://www.picard.ch/downloads/Hsp90interactors.pdf. 2017.

52.   Masel J. Evolutionary capacitance may be favored by natural selection. Genetics. 2005 Jul;170(3):1359-71.

53.   Masel J. Cryptic genetic variation is enriched for potential adaptations. Genetics. 2006 Mar;172(3):1985-91.

54.   Sangster TA, Lindquist S, Queitsch C. Under cover: causes, effects and implications of Hsp90-mediated genetic capacitance. Bioessays. 2004 Apr;26(4):348-62.

55.   Queitsch C. Thermotolerance, buffering of genetic variation and developmental stability - different aspects of chaperone function in the plant *Arabidopsis thaliana*. Chicago: University of Chicago; 2001.

56.   Mason GA, Lemus T, Queitsch C. The Mechanistic Underpinnings of an ago1-Mediated, Environmentally Dependent, and Stochastic Phenotype. Plant Physiol. 2016 Apr;170(4):2420-31.

57.   Lachowiec J, Queitsch C, Kliebenstein DJ. Molecular mechanisms governing differential robustness of development and environmental responses in plants. Ann Bot. 2016 Apr;117(5):795-809.

58.   Salathia N, Lee HN, Sangster TA, Morneau K, Landry CR, Schellenberg K, et al. Indel arrays: an affordable alternative for genotyping. Plant J. 2007 Aug;51(4):727-37.

59.   Debat V, Milton CC, Rutherford S, Klingenberg CP, Hoffmann AA. Hsp90 and the quantitative variation of wing shape in Drosophila melanogaster. Evolution. 2006 Dec;60(12):2529-38.

60.   Karras GI, Yi S, Sahni N, Fischer M, Xie J, Vidal M, et al. HSP90 Shapes the Consequences of Human Genetic Variation. Cell. [Twin Study]. 2017 Feb 23;168(5):856-66 e12.

61.   Sahni N, Yi S, Taipale M, Fuxman Bass JI, Coulombe-Huntington J, Yang F, et al. Widespread macromolecular interaction perturbations in human genetic disorders. Cell. 2015 Apr 23;161(3):647-60.

62.   Jarosz DF, Taipale M, Lindquist S. Protein homeostasis and the phenotypic manifestation of genetic diversity: principles and mechanisms. Annu Rev Genet. 2010;44:189-216.

63.   Dorrity MW, Cuperus JT, Carlisle JA, Fields S, Queitsch C. Balance between mutually exclusive traits shifted by variants of a yeast transcription factor. bioRxiv. 2018;doi: https://doi.org/10.1101/117911

64.   Citri A, Harari D, Shohat G, Ramakrishnan P, Gan J, Lavi S, et al. Hsp90 recognizes a common surface on client kinases. J Biol Chem. 2006 May 19;281(20):14361-9.

65.   Xu Y, Lindquist S. Heat-shock protein hsp90 governs the activity of pp60v-src kinase. Proceedings of the National Academy of Sciences of the United States of America. 1993 Aug 1;90(15):7074-8.

66.   Xu Y, Singer MA, Lindquist S. Maturation of the tyrosine kinase c-src as a kinase and as a substrate depends on the molecular chaperone Hsp90. Proceedings of the National Academy of Sciences of the United States of America. 1999 Jan 5;96(1):109-14.

67.   Stasinopoulos TC, Hangarter RP. Preventing photochemistry in culture media by long-pass light filters alters growth of cultured tissues. Plant Physiol. 1990 Aug;93(4):1365-9.

68.   Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009 Jul 15;25(14):1754-60.

69.   Clark RM, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, et al. Common sequence polymorphisms shaping genetic diversity in Arabidopsis thaliana. Science. 2007 Jul 20;317(5836):338-42.

Figure Legends

Figure 1: HSP90 perturbation results in greater penetrance of embryonic lethal mutations.

(**A**) Experimental workflow across generations. (**B**) Viable M2 seeds per silique after EMS

treatment. EMS-treated seeds were grown, yielding M1 plants bearing siliques with M2 seeds. A

single silique was collected from 50 individuals for each genotype and viable and dead seeds per silique were counted. *p<0.05, Student's unpaired t-test. Error bars are standard deviation. (**C**) Estimated number of embryonic lethal mutations per diploid genome (μ) after EMS mutagenesis. μ was calculated as $\mu=\ln(Y)-\ln(X)$, where Y equals the number of siliques collected from M1 plants within a genotype and X equals the number of siliques without dead seeds. *p<0.05, **p<1.0E-06, Poisson regression. Error bars are standard error calculated from a Poisson distribution. (D) Distribution of unique and likely EMS-derived SNVs in six Col-0 M2 and six HSP90 RNAi-C1 M2 individuals. Shades of red denote SNVs detected in HSP90 RNAi-C1 M2 plants, and shades of gray denote SNVs in Col-0 M2 plants. Centromeres were not considered (light blue). (**E**) Frequency of unique and likely EMS-derived SNVs in Col-0 M2 and HSP90-RNAi M2 individuals. Error bars represent 95% confidence intervals. p=0.1460 (not significant), Student's unpaired t-test.

Figure 2: Phenotype scoring index for M1 and M2 seedlings. (**A**) Phenotype categories for 16 aberrant morphological traits in early seedlings, ranging from wild-type-like (category 0) to strongly affected, being either an aborted seedling or showing four or more morphological defects (category 4). (**B**) Example M2 seedlings with common aberrant morphological phenotypes at varying levels of severity. Scores for phenotypic categories are indicated. Scale bars equal 1 mm.

Figure 3: HSP90 perturbation increases the frequency of severely aberrant phenotypes in the M2 generation. (**A**) Experimental workflow with examined seedling populations. (**B**) Phenotype frequencies in response to different concentrations of the HSP90 inhibitor geldanamycin (GdA) in wild-type Col-0 (shades of grey) and mutagenized M2 seedlings (shades of red). Y-axis, percent of seedlings with mutant phenotype scores > 3 (see **Figure 2**). The following numbers of seedlings was examined for each genotype and treatment: wild-type Col-0: DMSO n=731, 0.3 μM GdA n=143, 0.6 μM GdA n=319, 1.0 μM GdA n=243; mutagenized M2: DMSO n=1394, 0.3 μM GdA n=492, 0.6 μM GdA n=933, 1.0 μM GdA n=1488. *p-adj<0.05, **p-adj<1.0E-07, ***p-adj<1.0E-15, Fisher's exact test. See **S2 Table** for p-values. (**C**) Genetic perturbation of HSP90 reproduces the effects of pharmacological inhibition shown in **B**. Unmutagenized in light shades; mutagenized in darker shades. The following number of seedlings was examined for each genotype and treatment: wild-type Col-0 control, n=347; mutagenized Col-0 M2, n=817. Unmutagenized HSP90 RNAi-A1 control, n=314; mutagenized M2 HSP90 RNAi-A1, n=688. Unmutagenized HSP90 RNAi-C1 control, n=362; mutagenized M2 HSP90 RNAi-C1, n=632. *p-adj<0.05, **p-adj<1.0E-07, Fisher's exact test. Error bars

represent standard error as calculated by: SE = sqrt(p(100-p)/n), where p equals the percentage of seedlings with aberrant phenotypes and n equals the number of observations.

Figure 4: Aberrant morphological phenotypes are heritable and HSP90-responsive across generations in several mutagenized lines. (**A**) Experimental work flow with examined seedling populations. (**B**) Summary table of phenotype heritability in M3 lines. (**C**) Frequencies of the deformed aerial organ phenotype in wild-type Col-0 and mutagenized M3 populations in response to HSP90-reduced (0.5 μM GdA) or mock growth conditions (DMSO). Black lines represent the response of wild-type Col-0 to HSP90 perturbation in three biological replicates. Shades of red denote lines with 10% (5% for frequency of aborted seedlings) increase in phenotype frequency under HSP90-reduced conditions and shades of blue denote lines with 10% decrease in phenotype frequency under HSP90-reduced conditions (5% for frequency of aborted seedlings). Gray line color denotes non-responsive lines. The following number of seedlings was examined: n~144 for wild-type Col-0 and n~72 for each M3 line. *p-adj<0.05, Fisher's exact test. *p-adj<0.05, Fisher's exact test. (**D**) Heatmap of clustered odds ratios for GdA vs. DMSO (mock control) comparisons as calculated for wild-type Col-0 (green outline) and the mutagenized M3 lines across 16 early-seedling phenotypes (see **S2 Table** for all p-values and odds ratios (OR), **Figure 2**). Blue outline indicates M3 lines showing HSP90-responsive phenotypes commonly observed in Col-0 wild-type seedlings upon HSP90 perturbation albeit at lesser frequency (smaller OR) and severity. *p-adj<0.05, Fisher's exact test.

Figure 5: Bulk segregant analysis identifies candidate locus associated with HSP90-responsive trait. (**A**) Schematic of experimental work flow indicating provenance of examined

seedling populations. (**B**) Image of the selected M2 parent GM6, showing the deformed aerial organ phenotype that was successfully propagated to the M3 generation. Scale bar equals 5mm. (**C**) The GM6 M3 line showed increased frequency of seedlings with deformed aerial organs compared to wild-type Col-0. A GM6 M3 individual with deformed aerial organs was then backcrossed to wild-type Col-0. The resulting F2 population showed significantly increased frequency of the deformed organs trait when HSP90 was inhibited with GdA. Errors bars represent standard error of the mean across two biological replicates. *p-adj<0.05, **p-adj<1.0E-04, ***p-adj<1.0E-14, Fisher's exact test. (**D**) Bulk segregant analysis of GM6 F2 populations identifies a candidate locus associated with the HSP90-responsive trait deformed aerial organs. F2 individuals displaying this trait on HSP90-reduced conditions (n=94) were sequenced. Derived allele frequencies (y-axis) were estimated as the fraction of reads supporting the derived allele divided by the number of reads mapping to a given locus (see Materials and Methods). Dark gray box indicates the centromere and light gray box represents pericentromeric regions (chr1:12,000,000-18,000,000; chr2:1,000,000-7,500,000; chr3:11,000,00-17,000,000; chr4:1,800,000-7,500,000; chr5:10,000,000-17,000,000). Dashed red line represents derived allele frequency of 0.675. Dashed black line represents expected derived allele frequency of 0.5. The candidate locus contains one SNV (chr1:5,407,084 C=>T) encoding a missense mutation (proline to leucine) in the MYB family transcription factor AT1G15720.

Supporting Information

S1 Fig. Analysis pipeline for SNV calling in M2 plants. See Materials and Methods for details and S1 Table for results.

S2 Fig. HSP90 perturbation increases penetrance of new mutations in the M_1 generation.
(A) Schematic of experimental work flow. **(B)** Genetic perturbation of *HSP90* yields significantly increased frequency of aberrant phenotypes in M_1 HSP90 RNAi-C1-seedlings. Seeds of the indicated genetic backgrounds were treated with 0.2% EMS, plated on growth medium, seedlings were grown for ten days under standard growth conditions and assessed for 16 early-seedling traits (see **Figure 2**). The following number of seedlings was examined: wild-type Col-0 control n=213; mutagenized Col-0 M_1 n=948; un-mutagenized HSP90 RNAi-A1 control n=206; mutagenized HSP90 RNAi-A1 M_1 n=1009; un-mutagenized HSP90 RNAi-C1 control n=206; mutagenized HSP90 RNAi-C1 M_1 n=961. *p <0.05, Fisher's exact test. Error bars represent standard error as calculated by: SE = sqrt($p(100\text{-}p)/n$), where *p* equals the percentage of affected seedlings and *n* equals the number of observations (see **S2 Table** for all odds ratios and p-values). **(C)** Pharmacological inhibition of HSP90 reproduces the result observed in (**B**) for HSP90 RNAi-C1. Seeds, treated with either 0.2% or 0.4% EMS, were plated on media containing either DMSO or 1.0 μM GdA, grown for ten days under standard growth conditions and assessed for 16 early-seedling traits (see **Figure 2**). The following number of seedlings was examined: wild-type Col-0 control on DMSO n=144; 0.2% EMS-treated Col-0 on DMSO n=530; 0.4% EMS-treated Col-0 n=355 on DMSO; wild-type Col-0 control on 1.0 uM GdA n=36; 0.2% EMS-treated Col-0 on 1.0 uM GdA n=540; 0.4% EMS-treated Col-0 n=329. ***p-adj<1.0E-08, Fisher's exact test.

S3 Fig. Aberrant phenotypes are heritable and HSP90-responsive across generations in several mutagenized lines. Frequencies of 15 aberrant phenotypes in wild-type Col-0 and mutagenized M_3 lines in response to HSP90 perturbation (0.5 uM GdA) or mock conditions (DMSO). Black lines represent the response of wild-type Col-0 to HSP90 perturbation in three biological replicates. Shades of red denote lines with 10% (5% for frequency of aborted seedlings) increase in phenotype frequency under HSP90-reduced conditions and shades of blue denote lines with 10% decrease in phenotype frequency under HSP90-reduced conditions (5% for frequency of aborted seedlings). Gray line color denotes non-responsive lines. The following number of seedlings was examined per treatment: wild-type Col-0 n~144 in each of three replicate experiments; mutagenized M_3 lines n~72.

S4 Fig. Several M_3 lines propagated from M_2 parents with deformed aerial organs show higher frequency of this trait in response to HSP90 perturbation than Col-0 wild-type. Heatmap of clustered odds ratios for 0.5 μM GdA-treated wild-type Col-0 seedlings vs. 0.5 μM GdA-treated seedlings of indicated M_3 lines as calculated for each of the 16 early-seedling phenotypes. Black asterisk (*), p-adj<0.05, Fisher's exact test, significantly greater trait frequency in indicated M_3 lines compared to wild-type Col-0 seedlings in response to HSP90 perturbation.

S5 Fig. Bulk segregant analysis identifies candidate locus associated with HSP90-responsive trait. (A) Schematic of experimental work flow indicating provenance of examined seedling populations. **(B)** Image of the selected M_2 parent GM9, showing the deformed aerial organ

phenotype that was successfully propagated to the M_3 generation. Scale bar equals 5mm. **(C)** The GM9 M_3 line showed increased frequency of seedlings with deformed aerial organs compared to wild-type Col-0. A GM9 M_3 individual with deformed aerial organs was then backcrossed to Col-0. The resulting F_2 population showed significantly increased frequency of the deformed organs trait when HSP90 was inhibited with GdA. Errors bars represent standard error of the mean across two biological replicates. *p-adj<0.05, **p-adj<1.0E-04, ***p-adj<1.0E-14, Fisher's exact test. **(D)** Bulk segregant analysis of GM9 F_2 populations identifies a candidate locus associated with the HSP90-responsive trait deformed aerial organs. F_2 individuals displaying on HSP90-reduced conditions (n=117) were sequenced. Derived allele frequencies (y-axis) were estimated as the fraction of reads supporting the derived allele divided by the number of reads mapping to a given locus (see **Materials and Methods**). Dark gray box indicates the centromere and light gray box represents pericentromeric regions (chr1:12,000,000-18,000,000; chr2:1,000,000-7,500,000; chr3:11,000,00-17,000,000; chr4:1,800,000-7,500,000; chr5:10,000,000-17,000,000). Dashed red line represents derived allele frequency for the lowest observed derived allele frequency for the six candidate mutations (**S4 Table**). Dashed black line represents expected derived allele frequency of 0.5. The candidate locus contains six transition SNVs (**S4 Table**).

S6 Fig. Bulk segregant analysis data of backcrossed GM6 and GM9 F_2 populations. Allele frequency estimations at likely EMS-derived SNVs in GM6 **(A)** and GM9 **(B)** backcrossed F_2 populations. F_2 individuals displaying this trait on HSP90-reduced conditions (GDA) were sequenced. Derived allele frequency estimations across all five chromosomes are shown (bp; x-axis). Dark gray boxes indicate centromeres(1) and light gray boxes represent pericentromeric

regions (chr1:12,000,000-18,000,000; chr2:1,000,000-7,500,000; chr3:11,000,00-17,000,000; chr4:1,800,000-7,500,000; chr5:10,000,000-17,000,000). Derived allele frequencies (y-axis) were estimated as described in the **Materials and Methods**. Dashed red line represents derived allele frequency for the lowest observed derived allele frequency for the six candidate mutations. Dashed black line represents expected derived allele frequency of 0.5. The candidate locus contains six transition SNVs (**S4 Table**).

S7 Fig. Derived allele frequencies for 100 simulations. Simulation, described fully in the Materials and Methods section, assumes a given number of SNVs (n), recombining length of chromosome (L), and x-coverage (x), each based on actual data. The final simulation is shown in bold to illustrate typical levels of fluctuation in derived allele frequency for one simulation. Note that while derived allele frequency at simulated SNVs at times exceeds 0.65 (red dashed line), none of the 100 simulations contained five or more consecutive SNVs with this frequency.

S1 Table. Number of SNVs in mutated and unmutated plants.

S2 Table. All p-values.

S3 Table. Parental M2 phenotypes.

S4 Table. Candidate SNVs with elevated DAF.

(A)
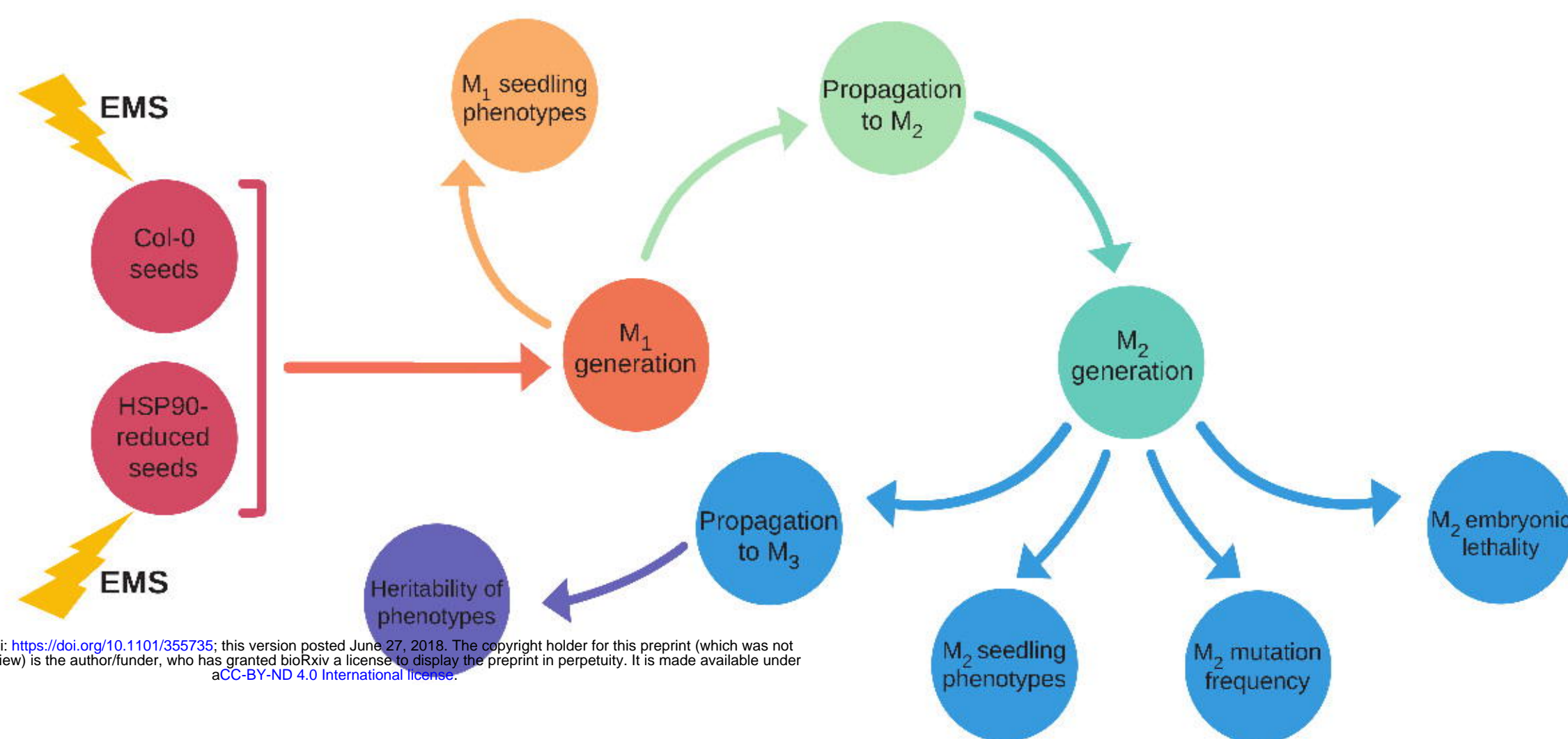
(B) M₂ seed viability

*Unpaired t-test: p<0.05
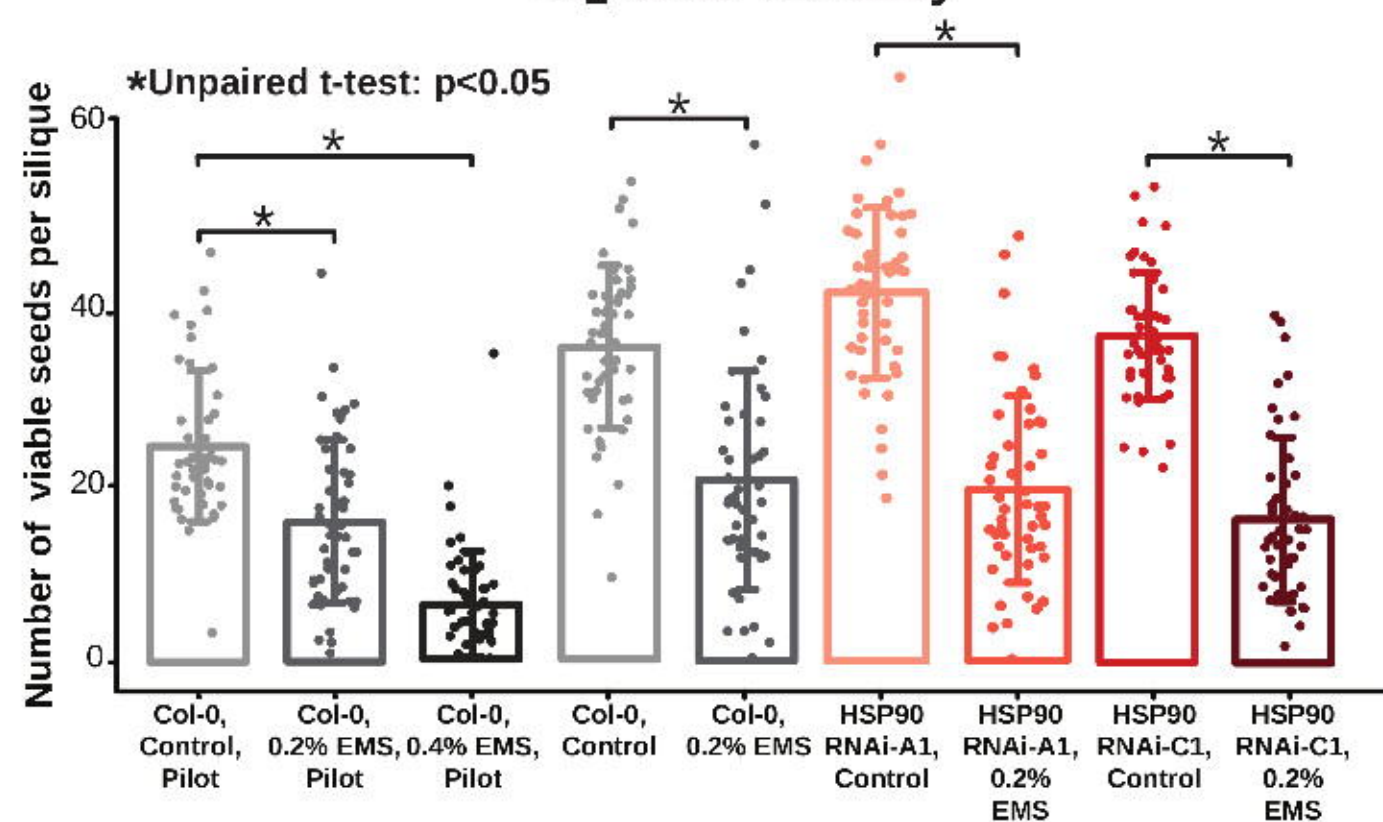
(C) Average number of embryonic lethal mutations

*glm: p<0.05
**glm: p<1.0E-06
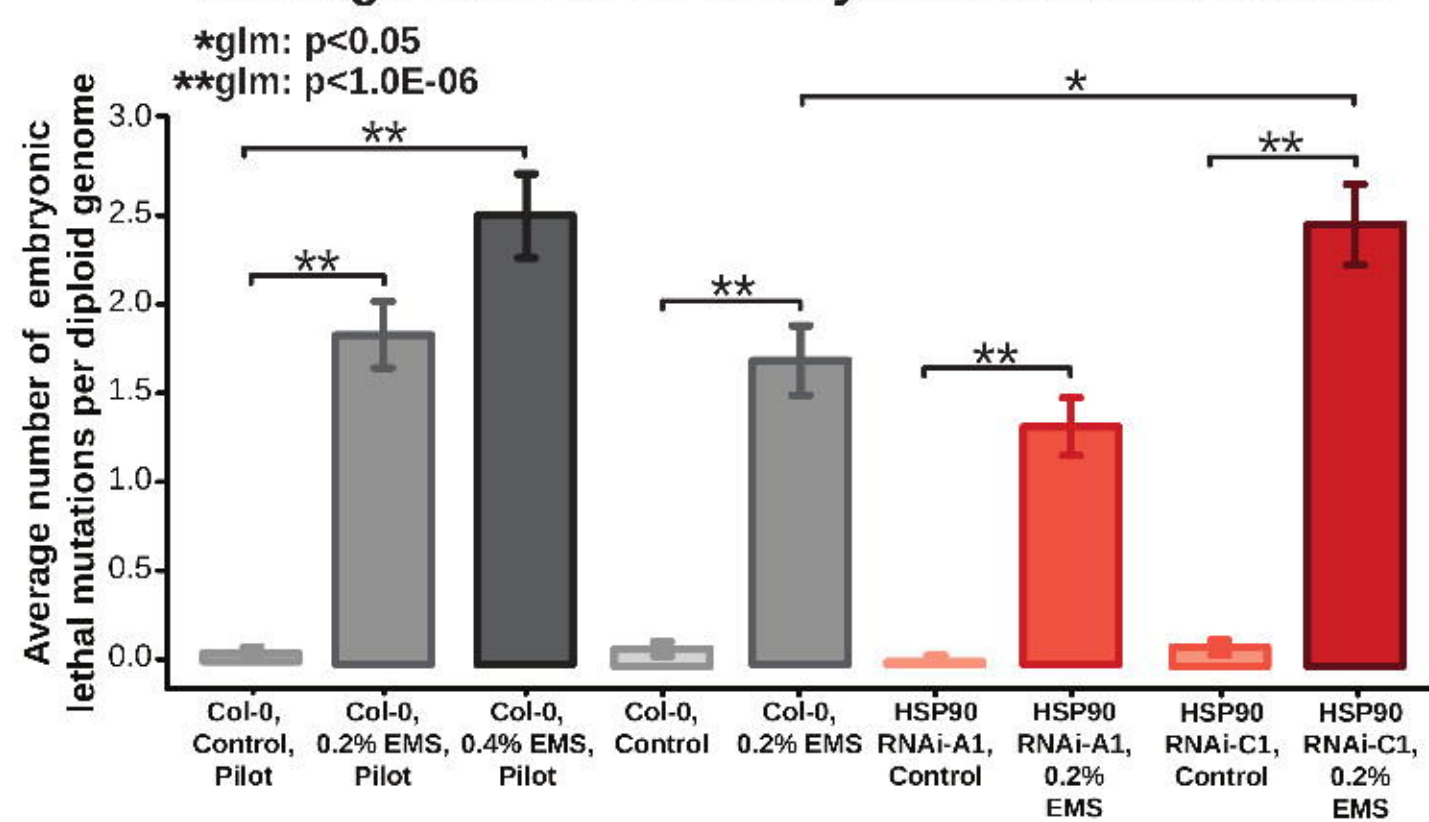
(D) Distribution of new SNVs

Col-0 M₂ individuals 1 2 3 4 5 6
HSP90 RNAi-C1 M₂ individuals 1 2 3 4 5 6

Chr 1
Chr 2
Chr 3
Chr 4
Chr 5
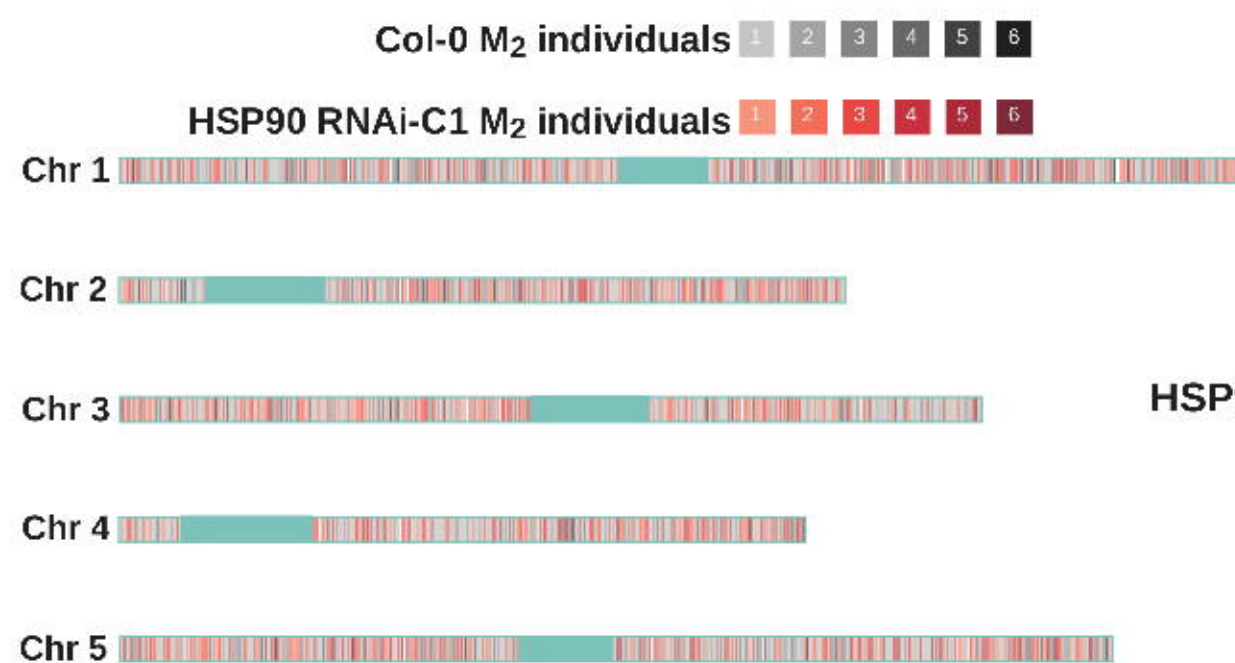
(E) Frequency of new SNVs
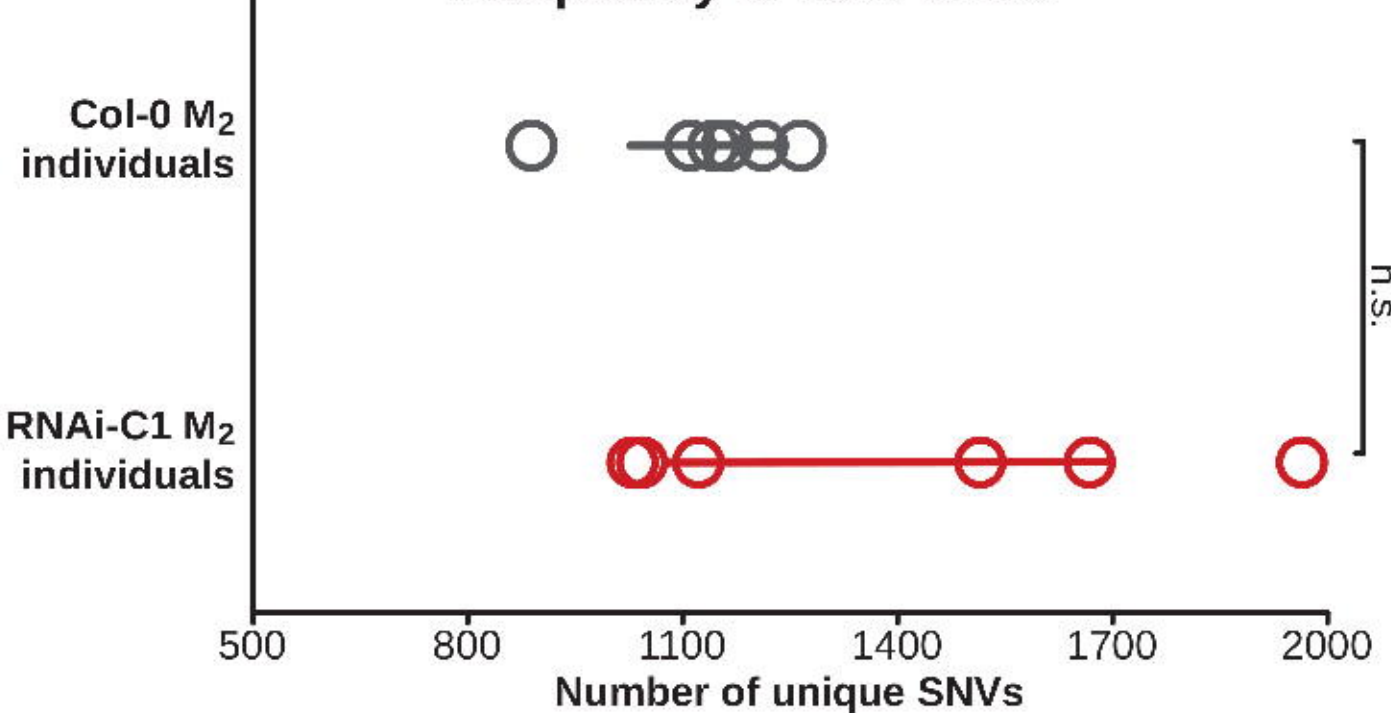
(A)

Category 0: No visible defects
- Cotyledons
- True leaves

Category 1: One defect

Rosette asymmetries | Organ asymmetries

Etiolated organs | Epinastic true leaves | Discolorations | Dwarf

Epinastic cotyledons | Hyponastic cotyledons | Lesions on cotyledons | Gravitropism defects
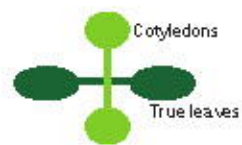
Category 2: Two defects, OR:
- Extra organs
- Missing organs
- Deformed organs
- Lesions on true leaves
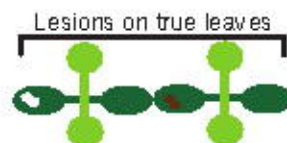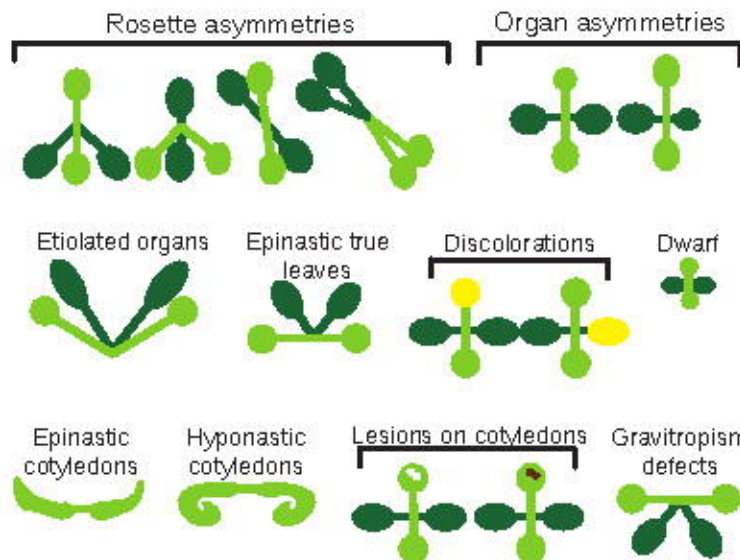
Category 3: Three defects, OR:
- Severe developmental delay

Category 4: Four or more defects, OR:
- Aborted seedling

(B)

Score 0 | Score 1 | Score 1
Score 2 | Score 2 | Score 2
Score 3 | Score 3 | Score 3
Score 4 | Score 4 | Score 4

(A) M₁ generation → Propagation to M₂ → M₂ generation → M₂ early-seedling phenotypes

(B) Col-0, M₂

Percent seedlings with mutant phenotype score ≥ 3

Fisher's exact test
***P<1.0E-15
**P<1.0E-07
*P<0.05

DMSO | 0.3 μM GdA | 0.6 μM GdA | 1.0 μM GdA

(C) Control, M₂

Percent seedlings with mutant phenotype score ≥ 3

Fisher's exact test:
**P<1.0E-07
*P<0.05

Col-0 | HSP90 RNAi-A1 | HSP90 RNAi-C1

(A)

M₁ generation → Propagation to M₂ → M₂ generation → Select mutant-like M₂ seedlings → Propagation to M₃ → Heritability of phenotypes in M₃ → Validation of phenotypes in F₂

(B)

Aberrant parental phenotype: GM6 M₂

Deformed aerial organs

(C)

Fisher's exact test:
*p<0.05
**p<1.0E-04
***p<1.0E-14

DMSO
0.5 µM GdA

Percent seedlings with deformed aerial organs

Col-0   GM6 M₃

Col-0   GM6 F₂

(D)
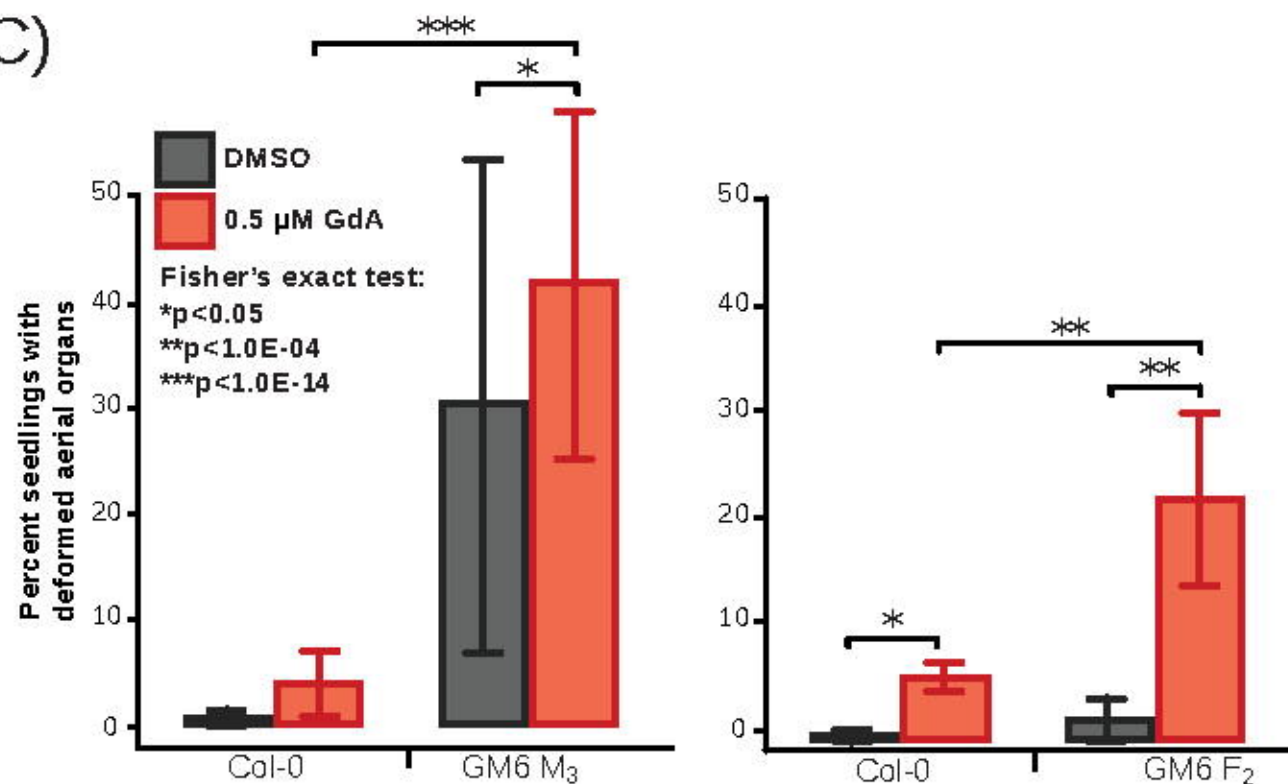
Derived allele frequency on Chromosome 1

Derived allele frequency

146 SNVs

Position on Chromosome 1